AWARD NUMBER: W81XWH-14-1-0596

TITLE: Biomarkers for Early Detection of Clinically Relevant Prostate Cancer. A Multi-Institutional Validation Trial

PRINCIPAL INVESTIGATOR: Jesse McKenney, MD

CONTRACTING ORGANIZATION: Cleveland Clinic Foundation
Cleveland, OH 44195

REPORT DATE: October 2016

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE | 2. REPORT TYPE | 3. DATES COVERED |
|---|---|---|
| October 2016 | Annual | 30 Sept 2015 – 29 Sept 2016 |

**4. TITLE AND SUBTITLE**

Biomarkers for Early Detection of Clinically Relevant Prostate Cancer. A Multi-Institutional Validation Trial

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
**W81XWH-14-1-0596**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
Jesse McKenney, MD

E-Mail: mckennj@ccf.org

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

CLEVELAND CLINIC FOUNDATION
9500 Euclid Ave.
Cleveland, OH 44195

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSOR/MONITOR'S ACRONYM(S)**

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**

Approved for Public Release; Distribution Unlimited

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

For men diagnosed with early stage prostate cancer a critical need exists for molecular assays that accurately distinguish aggressive prostate cancer from those cancers that will not cause harm if left untreated. In this project, we are assessing three different panels of established molecular biomarkers for their ability to distinguish aggressive cancers from indolent cancers. We have established agreements with three commercial companies to analyze their biomarker platforms in our multi-center, prospectively accrued prostate cancer active surveillance cohort – the Canary Prostate Active Surveillance Study (PASS). We are in the process of evaluating these three biomarker panels in tissue, blood, and urine samples with well annotated clinical and pathologic data collected as part of PASS. We are conducting rigorous statistical evaluation to demonstrate the utility and performance of biomarkers in clinical practice to predict aggressive disease. The accuracy of each biomarker for predicting short- and long-term progression will be characterized with time dependent receiver operating characteristic curves. The successful clinical validation of biomarkers that offer substantially improved predictive and prognostic accuracy should bring extraordinary potential to improve the care of prostate cancer patients.

**15. SUBJECT TERMS**
Prostate cancer, active surveillance, biomarkers, validation

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | USAMRMC |
| U | U | U | UU | 49 | 19b. TELEPHONE NUMBER *(include area code)* |

## Table of Contents

## 1. INTRODUCTION

Although prostate-specific antigen (PSA) testing and the resulting treatment of prostate cancer (PCa) is likely responsible for some of the 44% decrease in prostate cancer mortality witnessed in the United States since 1992, the detection of low risk tumors has increased. The majority of prostate cancers currently diagnosed are low risk tumors for which there is substantial evidence that the cancer will not cause harm if left untreated. However, enough uncertainty remains in accurately identifying which tumors will not cause harm to a patient that many low risk cancers are still treated, resulting in so-called overtreatment. To reduce this overtreatment, while still diagnosing aggressive high risk tumors early enough that they can be successfully treated, there is a critical need for molecular assays that accurately distinguish more aggressive disease from cancers that will not cause harm. The goal of this project is to perform rigorous clinical validation of established biomarkers in order to improve the accuracy of risk assessment and distinguish aggressive from indolent disease in men with apparently low-risk disease by standard clinical variables. We are evaluating multiple established and analytically validated quantitative molecular biomarkers to predict PCa progression in a multi-center active surveillance cohort with high-quality biospecimens. We aim to unlink the diagnosis of PCa with immediate treatment, thus addressing the overtreatment issue and economic, physical, and emotional burdens of PCa diagnoses. The results have promise to change the standard of care in the treatment of the majority of newly diagnosed PCa with near term impact due to the availability of the biomarkers and execution in an established, prospective cohort of men undergoing AS.

## 2. KEYWORDS

Prostate cancer; active surveillance; progression; aggressive disease; central pathology review; biomarkers; prediction models; PCA3; TMPRSS2:ERG; kallikreins; 4Kscore; OncotypeDX;

## 3. ACCOMPLISHMENTS

**What were the major goals and objectives of the project?**

We hypothesize that biomarkers of disease aggressiveness and prognosis can be interrogated in low risk prostate cancer (PCa) and that these biomarkers will better detect clinically relevant PCa in asymptomatic patients, thus distinguishing aggressive from indolent disease and immediately impacting both the initial choice of therapy and decision-making during AS. The objective of the study is to utilize analytically validated assays that take into account tumor heterogeneity to measure biomarkers in specimens that were collected in a non-invasive manner.

The major goals of the project, as stated in the scope of work, are:

1. Collection of specimens and clinical data. (Coordinated by FHCRC)

Milestone 1. Completion of a minimum of three years of follow-up with high-quality data and specimen collection. Due: 12/30/2016

2. Analysis of scientific aim 1: Validate a panel of tissue-based biomarkers to determine the presence of or progression to aggressive disease. (Lead site: FHCRC)

Milestone 2. Execute collaboration agreement with GHI. Due 12/30/2014 COMPLETED.

**Milestone 3. Tissue blocks identified for analysis. Due: 12/30/2015 COMPLETED**

Milestone 4. Oncotype DX validation complete in PASS cohort. Due 12/30/2016

Milestone 5. Manuscript submission of Oncotype DX validation. Due 9/30/2017

3. Analysis of scientific aim 2: Evaluate a panel of four-kallikrein plasma-based markers to determine the presence of or progression to clinically relevant prostate cancer. (Lead site: FHCRC)

Milestone 6. Execute collaboration agreement with OPKO. Due 3/30/2015 COMPLETED.

**Milestone 7. Plasma samples identified for analysis. Due 12/30/2015 COMPLETED**

**Milestone 8. OPKO 4KScore validation complete in PASS cohort. Due 9/30/2016 COMPLETED**

Milestone 9. Manuscript submission of 4KScore validation. Due 9/30/2017

4. Analysis of scientific Aim 3: Confirm the ability of PCA3 mRNA concentrations in urine, alone or in combination with TMPRSS2:ERG mRNA. (Lead site: FHCRC)

Milestone 10. Urine specimens identified for analysis. Due 12/30/2014 COMPLETED

**Milestone 11. PCA3 and TMPRSS2:ERG validation complete in PASS cohort. Due 12/30/2015 COMPLETED**

Milestone 12. Manuscript submission of PCA3 and TMPRSS2:ERG validation. Due 9/30/2017

5. Central pathology review of PASS biopsy and RP slides. (Lead site: CCF)
   Milestone 13. Completion of Central Pathology Review for biopsy-driven endpoints. Due: 12/30/2016

6. Translation of biomarkers into clinical practice. (Lead sites: FHCRC and CCF)
   Milestone 14. Construction of integrated model of biomarkers for the prediction of progression in the PASS cohort. Due 9/30/2017
   Milestone 15. Manuscript submission of integrated model for prediction of progression. Due 9/30/2017

**What was accomplished under these goals?**

*Task 1: Collection of specimens and clinical data. (Coordinated by FHCRC)*

Collection of follow-up data and longitudinal specimens in the PASS cohort is essential to adequately power our funded biomarker analyses. To date, PASS has enrolled 1,379 eligible patients at nine clinical sites. We have been highly successful in following participants to obtain outcomes measures, with a median cohort follow-up of 4.1 years (25th and 75th percentiles: 2.2, 6.0 years). Currently, all of the first 1000 participants enrolled in PASS, which are the subject of this specific research proposal, have at least three years of follow-up. In the past year, we have conducted site visits to three clinical sites this year (Beth Israel Deaconess Medical Center, University of Washington and University of Texas Health Sciences Center, San Antonio) to ensure adherence to the protocol and the coordinating center based at the Fred Hutchinson Cancer Center continues to provide data QA and QC.

*Task 2: Analysis of scientific Aim 1: Validate a panel of tissue-based biomarkers to determine the presence of or progression to aggressive disease. (Lead site: FHCRC)*

To date, we have collected FFPE tissue blocks from the diagnostic biopsies of 577 PASS participants at the 9 different PASS sites and the local clinics associated with each site. The pathology reports for each case have been collected and the pathology data in the PASS database has been reviewed and checked for quality assurance. The blocks have been labeled with a PASS ID and sent to Genomic Health, Inc. for processing.

*Task 3: Analysis of scientific Aim 2: Evaluate a panel of four-kallikrein plasma-based markers to determine the presence of or progression to clinically relevant prostate cancer. (Lead site: FHCRC)*

We have collaborated with OPKO to assay a panel of four kallikreins (total PSA (tPSA), free PSA (fPSA), intact PSA (iPSA), and human kallikrein 2 (hK2)). Statistical models were developed to predict reclassification from Gleason 6 cancer to Gleason 7 or greater. The analysis plan was determined before specimens were selected for the study, and included

breaking the data/specimens into training and testing cohorts, using a 2/3 to 1/3 split. The models included clinical information and either the 4Kpanel or serum PSA. We used Receiver Operating Characteristic (ROC) curve analyses and area under the curve (AUC) to assess discriminatory capacity and decision curve analysis (DCA) to report clinical net benefit.

Significant predictors for reclassification were 4Kpanel (OR=1.54 [1.31,1.81]) or PSA (OR=2.11 [1.53,2.91]), ≥20% cores positive (OR=2.10 [1.33,3.32]), ≥2 prior negative biopsies (OR=0.19 [0.04,0.85]), prostate volume (OR=0.47 [0.31,0.70]), BMI (OR=1.09 [1.04,1.14]). ROC curve analysis comparing 4Kpanel and base models indicated that the 4Kpanel improved accuracy for predicting reclassification (AUC 0.78 versus 0.74) in the first surveillance biopsy. Both models performed comparably for prediction of reclassification in subsequent biopsies (AUC=0.75 versus 0.76). In DCA, both models showed higher net benefit compared to biopsy-all and biopsy-none strategies.

Conclusions: The 4Kpanel provided incremental value over routine clinical information in predicting high-grade cancer in the first biopsy after diagnosis. The 4Kpanel did not add predictive value to the base model at subsequent surveillance biopsies.

These results were presented at the 2016 Meeting of the American Urological Association (AUA) and have been submitted for publication. Further details are in the attached manuscript that is currently under a third round of review at European Urology.

*Task 4: Analysis of Specific Aim 3: Confirm the ability of PCA3 mRNA concentrations in urine, alone or in combination with TMPRSS2:ERG mRNA to predict the presence of or development to clinically relevant prostate cancer. (Lead site: FHCRC)*

PCA3 and the TMPRSS2:ERG fusion are prostate cancer-specific biomarkers that hold promise for stratifying risk in the setting of AS. Hologic Gen-Probe's assay to quantitate urine PCA3 transcripts in post-digital rectal exam (DRE) urine is FDA-approved for men with a previous negative biopsy, given peer reviewed evidence that it can reduce unnecessary prostate biopsies. We aim to confirm the ability of the PCA3 and TMPRSS2:ERG assays to predict aggressive prostate cancer in the entire PASS cohort. To this end, we have collaborated with Hologic Gen-Probe to assay 2,926 urine specimens from 1,107 PASS participants. Analyses are underway and we expect to publish results in year three of this project. Included with this report are preliminary analyses of the urine marker data completed during years one and two of funding.

In the analyses described here, we evaluated PASS participants in which urine was collected prior to the first surveillance biopsy (which is sometimes called the confirmatory biopsy), had a Gleason score ≤ 3+4, and a ratio of positive to total biopsy cores < 34%. The endpoint was any grade or volume reclassification, where volume reclassification is defined as an increase in the ratio of biopsy cores containing cancer to total cores collected from <34% to

34% or greater. We used the full set of 552 samples that met the inclusion criteria to build clinical models with and without the PCA3 and TMPRSS2:ERG (T2:ERG) markers. Clinical predictors considered in modeling included serum PSA (logged), age, body mass index (BMI), race (African American or other), clinical T-stage, diagnostic Gleason score, cores ratio from diagnostic biopsy (ratio of biopsy cores containing cancer to total cores; logged), months since diagnosis, prostate size (logged), 5ARI use, family history of prostate cancer, smoking status (current, former, never), and study site. Logistic regression was used to fit the models, with variable selection occurring using backwards selection procedures based on a p-value cutoff of 0.05. Only PSA, cores ratio, and prostate size remained in the base clinical model as significant predictions for reclassification. In the model with urine biomarkers, PCA3 and T2:ERG were forced to remain in the model.

Our first goal was to use the urine biomarkers “paired” with the first surveillance biopsy to evaluate if the urine markers improved prediction of adverse reclassification in the first surveillance biopsy. The baseline characteristics of the cohort are presented in Table 1. One hundred thirty participants (24%) reclassified at the first biopsy. PSA, prostate size, ratio of positive to total biopsy cores in the diagnostic biopsy, and PCA3 were significantly different between participants that reclassified and those that did not. Neither age nor T2:ERG were significantly different between the groups. As in all of our analyses, we asked if a model including the urine markers improves the prediction of outcome relative to a model with commonly available clinical variables alone. As shown in Table 2 below, in univariate analysis, PCA3 is significantly associated with reclassification (OR = 1.6 (1.2, 1.9), p-value 0.0001), as are PSA, cores ratio, and prostate size. In multivariate analysis, PCA3 is still significantly associated with reclassification in the first surveillance biopsy (OR = 1.3 (1.0, 1.7), p-value 0.02), as are PSA, cores ratio, and prostate size. T2:ERG is not associated with reclassification in either univariate or multivariate analysis. ROC curve analysis (Figure 1) comparing the model with only clinical variables to the models with clinical variables plus either PCA3, T2:ERG, or both markers, shows a very slight improvement in AUC upon the addition of PCA3 but not T2:ERG.

Our second goal was to evaluate if the urine biomarkers improved prediction of time to adverse reclassification in later surveillance biopsies. We used Cox Proportional Hazard modeling to evaluate whether baseline urine markers improved prediction of the time to adverse reclassification over clinical variables alone. The same cohort was used as described above, except that participants were excluded if the reclassification occurred on the same day as urine specimen collection. As shown in Table 3, both PSA and prostate size were significantly associated with reclassification, but in univariate and multivariate models, neither PCA3 or T2:ERG were associated with reclassification. ROC curve analysis comparing the model with only clinical variables to the models with clinical variables plus either PCA3,

T2:ERG, or both markers, for both two-year prediction and three-year prediction of adverse reclassification is shown in Figure 2. The AUC with 95% bootstrap CI for the clinical model with no urine biomarkers were 0.81 (0.74 – 0.88) for two-year prediction, and 0.77 (0.68 – 0.84) for three-year prediction. There was no substantial improvement when either PCA3 or T2:ERG were added to the models.

**Table 1.** Characteristics of PASS participants prior to first surveillance biopsy

| Variable | All Participants, n=552 Median [IQR] | Reclassifiers, n=130 Median [IQR] | Non-Reclassifiers, n=422 Median [IQR] | p-value* |
|---|---|---|---|---|
| **PSA** | 4.7 [2.9, 6.4] | 4.9 [3.7, 7.0] | 4.4 [2.7, 6.4] | 0.01 |
| **Prostate size** | 40.3 [30.0, 55.1] | 35.1 [25.4, 47.6] | 42.2 [31.6, 59.5] | <0.001 |
| **Cores ratio** | 8.3 [8.3, 16.7] | 16.7 [8.3, 24.5] | 8.3 [8.3, 16.7] | <0.001 |
| **Age at Dx** | 63 [58, 67] | 63 [58, 69] | 63 [58, 67] | 0.29 |
| **PCA3^** | 32 [18, 61] | 39.5 [24, 89] | 30 [16, 57] | <0.001 |
| **T2:ERG ^** | 14 [2, 57] | 27 [1, 82.5] | 13 [2, 53] | 0.24 |

^ Logged in models. Exponentiated for interpretability of median [IQR] in table.
* P-value from 2 sample t-test comparing reclassifiers to non-reclassifiers

**Table 2.** Model Fit Using 552 Paired PASS Urine Specimens to Predict Adverse Reclassification in Subsequent Biopsy

| Variable* | Univariate | | Multivariate | |
|---|---|---|---|---|
| | OR (95% CI)^ | p-value | OR (95% CI)^ | p-value |
| **PSA** | 1.5 (1.1, 2.0) | 0.01 | 1.8 (1.3, 2.6) | 0.001 |
| **Cores Ratio** | 4.0 (2.6, 6.3) | <0.001 | 3.4 (2.1, 5.4) | <0.001 |
| **Prostate Size** | 0.3 (0.2, 0.5) | <0.001 | 0.3 (0.1, 0.4) | <0.001 |
| **PCA3** | 1.6 (1.2, 1.9) | <0.001 | 1.3 (1.0, 1.7) | 0.02 |
| **T2:ERG** | 1.1 (1.0, 1.2) | 0.21 | 1.0 (0.9, 1.2) | 0.52 |

*all variables logged
^ Odds Ratio (OR) and 95% Confidence Interval (CI) correspond to 1 unit increase
Models with clinical variables, clinical variables + PCA3, clinical variables + T2:ERG give similar results

**Table 3.** Time to Adverse Reclassification in 405 PASS Participants with Urine Specimen Collected On or Before First Study Biopsy

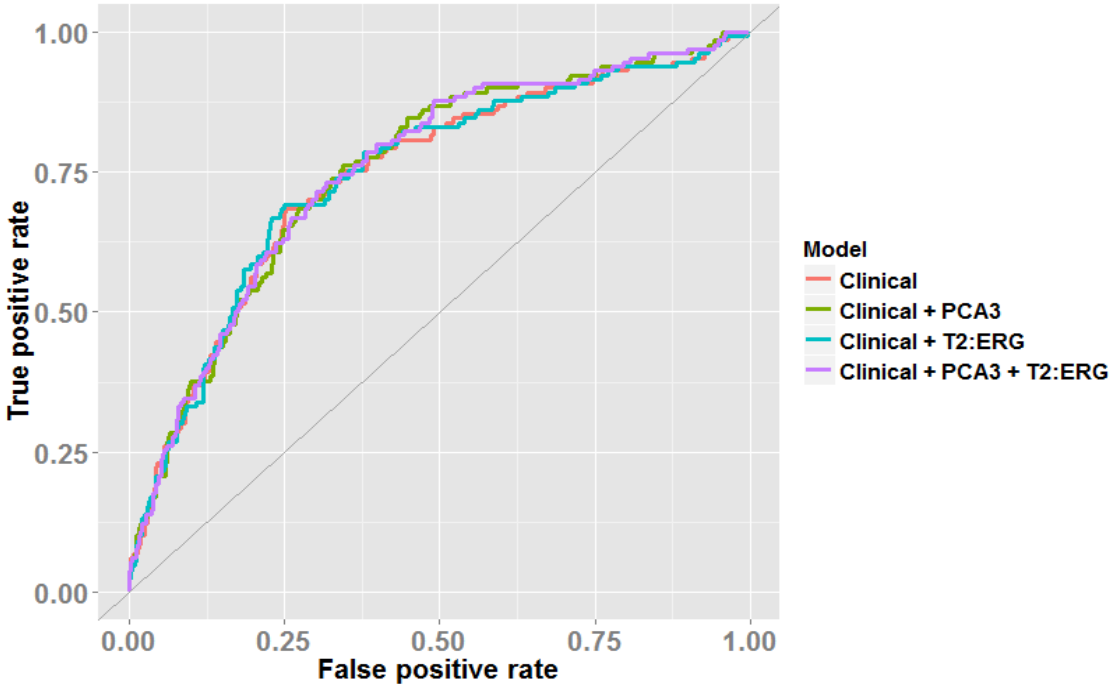
| Variable* | Univariate | | Multivariate | |
|---|---|---|---|---|
| | HR (95% CI)^ | p-value | HR (95% CI)^ | p-value |
| **PSA** | 2.1 (1.2, 3.5) | 0.006 | 2.9 (1.7, 4.9) | <0.001 |
| **Prostate Size** | 0.4 (0.2, 0.8) | 0.009 | 0.2 (0.1, 0.5) | <0.001 |
| **PCA3** | 1.3 (1.0, 1.7) | 0.09 | 1.2 (0.9, 1.6) | 0.32 |
| **T2:ERG** | 1.0 (0.9, 1.2) | 0.75 | 1.0 (0.9, 1.2) | 0.63 |

* Logged

^ Hazard Ratio (HR) and 95% Confidence Interval (CI) based on 1 unit increase

---

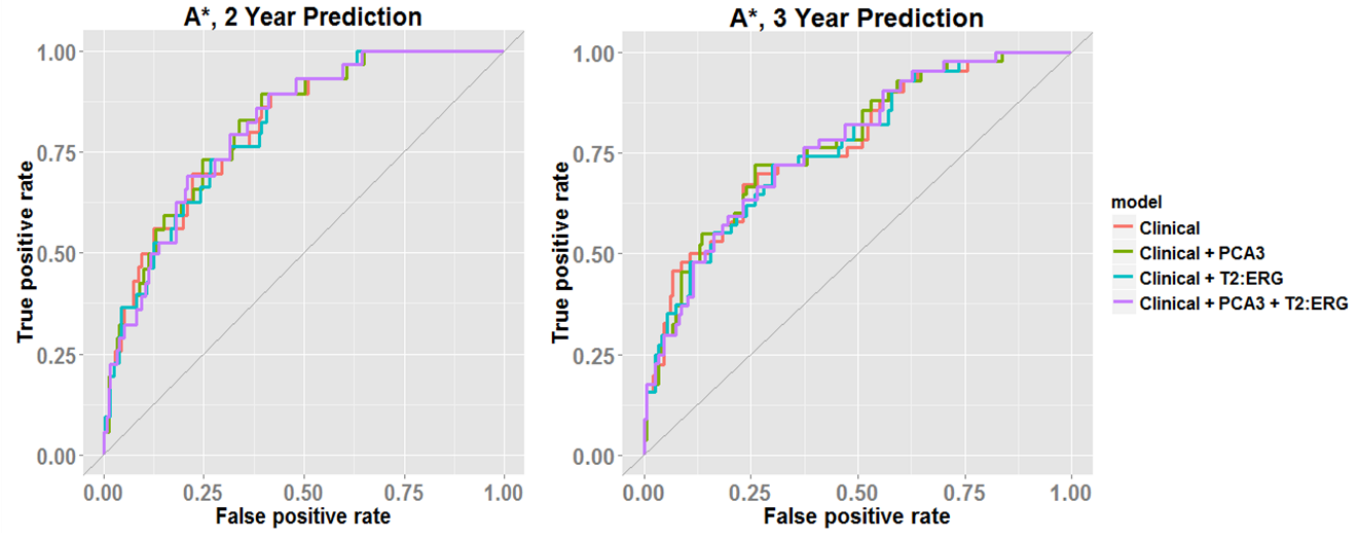
**Figure 1.** Prediction of Adverse Reclassification in 552 Paired PASS Urine Samples




| Model | AUC (95% CI) |
|---|---|
| Clinical variables alone (No urine) | 0.741 (0.687 – 0.790) |
| Clinical + PCA3 | 0.751 (0.699 – 0.796) |
| Clinical + T2:ERG | 0.742 (0.688 – 0.794) |
| Clinical + PCA3 + T2:ERG | 0.752 (0.701 – 0.799) |

---

**Figure 2.** Prediction of Time to Adverse Reclassification Analysis Using 405 PASS Participants with Urine Specimen Collected On or Before First Study Biopsy




| Model | AUC (95% Bootstrap CI) | |
|---|---|---|
| | 2 year | 3 year |
| Clinical variables alone (No urine) | 0.81 (0.74 – 0.88) | 0.77 (0.68 – 0.84) |
| Clinical + PCA3 | 0.82 (0.74 – 0.88) | 0.78 (0.70 – 0.85) |
| Clinical + T2:ERG | 0.81 (0.74 – 0.87) | 0.77 (0.68 – 0.84) |
| Clinical + PCA3 + T2:ERG | 0.81 (0.73 – 0.87) | 0.77 (0.70 – 0.85) |

*Task 5: Central Pathology Review (Lead Site: Cleveland Clinic)*

The purpose of central pathology review is to standardize endpoints for analyses of biomarkers. In the first two years of funding, we have developed a customized pathology review system, in which primary and secondary pathology reviewers can access scanned images and record key data from each slide. To date, we have scanned 1,025 slides from 661 diagnostic biopsies. All biopsies are reviewed by the primary pathologist as well as a secondary reviewer. All data recorded by the primary and secondary reviewers are reviewed for consistency and if results are discrepant, a consensus review is conducted to resolve. In an early analysis of scoring, we evaluated slides from 131 unique diagnostic biopsies, collected from five different PASS study sites. In this small subset, 71% of cases were reviewed concordantly by study pathologists and the original pathologist. The 29% discordant reviews highlight the need for a centralized review of cases to obtain accurate data, as Gleason is used as an endpoint in many biomarker studies.

**Table 4.** Analysis of Concordance in Central Pathology Review of 131 PASS Specimens

| Original Path | | Total Senarios | | | | |
|---|---|---|---|---|---|---|
| Gleason | Total Cases | Total Agreement | CR Agreement | Orig & 1° Agree | Orig & 2° Agree | Total Disagreement |
| 3 + 3 | 120 | 90 (76) | 12 (10) | 11 (9) | 4 (3) | 3 (2) |
| 3 + 4 | 8 | 2 (25) | 1 (13) | 0 | 5 (63) | 0 |
| 4 + 3 | 3 | 0 | 3 (67) | 0 | 0 | 0 |
| **TOTAL** | **131** | **92 (71)** | **16 (12)** | **11 (9)** | **9 (7)** | **3 (2)** |

**What opportunities for training and professional development did the project provide?**
Nothing to report. This grant does not provide for training or professional development activities.

**How were the results disseminated to communities of interest?**
Results are being disseminated through presentations at national meetings and through publication.

**What do you plan to do during the next reporting period to accomplish the goals and objectives?**
Our plans for the next year of funding are as follows:

- We will continue collecting follow-up data on the 1,000 PASS Study participants.
- We will complete tissue specimen acquisition for the analysis of the Oncotype Dx GPS score. We anticipate that we will lock clinical data and perform analysis validating the use of GPS in active surveillance.
- We will perform time-dependent analyses evaluating the performance of the 4Kpanel in predicting three-year and five-year reclassification.
- We will evaluate the performance of urine biomarkers in predicting adverse reclassification in the second and third surveillance biopsies.
- We have sent specimens to Hologic/Gen-Probe for assay to evaluate a secondary outcome of adverse pathology at time of surgery after a period of active surveillance. Once we have these results we will complete our analysis of urinary PCA3 and TMPRSS2:ERG in our active surveillance cohort.
- We will perform central review on follow-up surveillance biopsies and prostatectomy cases.

4**. IMPACT**

**What was the impact on the development of the principal discipline(s) of the project?**

We anticipate that the successful clinical validation of biomarkers that offer substantially improved predictive and prognostic accuracy would bring extraordinary potential to improve the care of PCa patients. Specifically, those men with clinically low-risk tumors that can be confirmed as truly low-risk with greater accuracy could be spared the cost and quality-of-life impact of invasive diagnostic and therapeutic procedures. Conversely, those men with apparent low-risk disease who in fact harbor higher-risk tumors or have the potential to develop lethal disease will be identified, thus avoiding under-treatment. Such a paradigm shift in PCa care would yield near-term changes in the PCa treatment landscape, greatly improving the cost-benefit calculations for population-level PCa screening efforts and reducing the overtreatment of disease.

**What was the impact on other disciplines?**

Nothing to report in this period, although we expect that statistical techniques being developed will be utilized to evaluate biomarker performance in many diseases other than prostate cancer.

**What was the impact on technology transfer?**

This project involves evaluation and validation of commercial biomarker panels that have not previously been used in the active surveillance setting. While we do not expect a direct impact on technology transfer, there should be a large impact on the commercial use of the molecular diagnostics.

**What was the impact on society beyond science and technology?**

Successful execution of this project should transform the clinical management of prostate cancer in several ways. First, if patients and their physicians have a reliable and valid estimate of the risks of disease progression and harm, then more might opt for surveillance, thereby reducing the risks of overtreatment and its attendant substantial costs and morbidity. Such improved accuracy would allow men to be selected more appropriately and with greater confidence for surveillance rather than immediate treatment. Second, a proportion of men initially choosing active surveillance eventually opt for primary curative treatment even with no objective measures of clinical progression, presumably due to patient/provider anxiety. Increasing patient and provider confidence in risk assessments would presumably lead to increased adherence to active surveillance, further decreasing overtreatment. Third, a marker panel with high accuracy for progression on active surveillance will influence the regimen of clinical re-assessment, such that those men with particularly low-risk disease might be eligible

for a less intensive surveillance protocol with fewer repeated prostate biopsies, reducing the use of the most invasive, and risky, component of a typical surveillance regimen. Fourth, the proposed markers might also facilitate treatment planning for men not currently on surveillance. For example, a man with apparently low-risk disease but a significantly adverse biomarker panel would have an increased risk of occult high grade disease and perhaps should undergo staging lymphadenectomy at time of prostatectomy, a procedure which might not routinely be performed for low risk disease. Lastly, the public health impact of a validated biomarker panel will be substantial, as the costs of initial curative therapy for prostate cancer accounts for $2-3 billion annually. Approximately half of the new diagnoses are low risk cancers and candidates for active surveillance, and accurate determination of who may benefit from curative therapy, while sparing the majority, would have immediate economic impact.

## 5. CHANGES / PROBLEMS

**Changes in approach and reasons for change**
Nothing to report.

**Actual or anticipated problems or delays and actions or plans to resolve them**
Nothing to report.

**Changes that had a significant impact on expenditures**
Nothing to report.

**Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents:**

- **Significant changes in use or care of human subjects:** No significant changes in the use or care of human subjects. The Fred Hutchinson Cancer Research Center has approved the study activities through 5/29/2017 under IR file number 8271. A continuing review was submitted to HRPO (Log Number A-18320) on 05/27/2016 and receipt was acknowledged on 07/01/2016.
- **Significant changes in use or care of vertebrate animals:** Nothing to report.
- **Significant changes in use of biohazards and/or select agents:** Nothing to report.

## 6. PRODUCTS

### Publications, Conference Papers, and Presentations

#### Journal publications

Daniel W. Lin, Lisa F. Newcomb, Marshall D. Brown, Daniel D. Sjoberg, Yan Dong, James D. Brooks, Peter R. Carroll, Matthew Cooperberg, Atreya Dash, William J. Ellis, Michael Fabrizio, Martin E. Gleave, Todd M. Morgan, Peter S. Nelson, Ian M. Thompson, Andrew Wagner, and Yingye Zheng for the Canary Prostate Active Surveillance Study Investigators. “Evaluating the four kallikrein panel of the 4Kscore for prediction of high-grade prostate cancer in 2 men in the Canary Prostate Active Surveillance Study (PASS).” European Urology. Under review. Acknowledgement of federal support: Yes.

#### Books or other non-periodical, one-time publications

Nothing to report.

#### Other publications, conference papers, and presentations

Lin D, Brown M, Newcomb L, Sjoberg D, Brooks J, Carroll P, Dash A, Fabrizio M, Gleave M, Morgan T, Nelson P, Thompson I, Zheng Y.  PD08-02: “Evaluating the four kallikrein panel of the 4KScore for prediction of high-grade prostate cancer in men in the Canary Prostate Active Surveillance Study (PASS).” Annual Meeting of the American Urological Association; 2016 May 6-10, San Diego, CA.

Newcomb L. “Evaluating urinary PCA3 and TMPRSS2:ERG for prediction of adverse biopsy reclassification in men in the Canary Prostate Active Surveillance Study (PASS).” Presentation at the Multi-Institutional Prostate Cancer SPORE Program Retreat, 2016 March 13-15, Fort Lauderdale, FL.

#### Website(s) or other Internet site(s)

Nothing to report.

#### Technologies or techniques

Nothing to report.

#### Inventions, patent applications, and/or licenses

Nothing to report.

**Other Products**

As part of this project we continue to maintain a large biospecimen repository with associated clinical and demographic data, which serves as a rich resource for the scientific community. In the coming years of this award we anticipate scientific results, validated diagnostics, and prediction models that should make an impact on the clinical management of patients with prostate cancer.

## 7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

**What individuals have worked on the project?**

| | |
|---|---|
| Name: | Daniel Lin, MD |
| Project Role: | Principal Investigator |
| Researcher Identifier (e.g. ORCID ID): | ORCID: 0000-0002-2135-1534 |
| Nearest person month worked: | 2 person months |
| Contribution to Project: | As Principal Investigator, Dr. Lin oversees the execution of the project, including interactions with industry collaborators and the FDA. He directs overall scientific activities including data collection, interpretation, and manuscript preparation. Dr. Lin takes a central role in the analysis of all data from the project, collaborating with the other investigators on manuscript preparations. |
| Funding Support: | N/A |

| | |
|---|---|
| Name: | Jesse McKenney, MD |
| Project Role: | Principal Investigator of Partner Award |
| Researcher Identifier (e.g. ORCID ID): | N/A |
| Nearest person month worked: | 2 person months |
| Contribution to Project: | Dr. McKenney is the lead pathologist for this project, overseeing all aspects of the central pathology review. He has worked on development of the Centralized Pathology Review system, and leads the group of study pathologists who review all endpoints for PASS participants. He ensures that pathologic review is timely and follows project guidelines. |
| Funding Support: | N/A |

| | |
|---|---|
| Name: | Hilary Boyer |
| Project Role: | Research Scientist |
| Researcher Identifier (e.g. ORCID ID): | N/A |
| Nearest person month worked: | 3 person months |
| Contribution to Project: | Ms. Boyer works under the direction of Dr. Newcomb to receive, annotate, and track PASS specimens from the Central Repository. Ms. Boyer is responsible for pulling, tracking, and documenting specimens sent to collaborating sites and coordinates all shipping activities. She also assists in specimen and clinical data QA and QC, in monitoring study progress, and in preparing reports for study investigators. |
| Funding Support: | N/A |

| | |
|---|---|
| Name: | Anna Faino, MS |
| Project Role: | Statistical Research Associate |
| Researcher Identifier (e.g. ORCID ID): | N/A |
| Nearest person month worked: | 7 person months |
| Contribution to Project: | Ms. Faino works under the supervision of Dr. Zheng and is responsible for the extensive data analysis involved in this project. She participates in study consultation with project investigators and the data operations group on data and database forms. Under Dr. Zheng's supervision she performs data analyses, data interpretation and manuscript preparation. |
| Funding Support: | N/A |

| | |
|---|---|
| Name: | Suzanne Kolb, MPH |
| Project Role: | Project Coordinator |
| Researcher Identifier (e.g. ORCID ID): | ORCID: 0000-0002-6443-644X |
| Nearest person month worked: | 6 person months |
| Contribution to Project: | Ms. Kolb works under the direction of Drs. Lin and Newcomb to fulfill daily fiscal and administrative functions of the program. She monitors subaward budgets and provides logistical support. Ms. Kolb works closely with the PASS Deputy Director to maintain IRB files, material transfer agreements, and other regulatory documents as well as tracking project timelines and deliverables. |
| Funding Support: | N/A |

| | |
|---|---|
| Name: | Lisa Newcomb, PhD |
| Project Role: | Deputy Director |
| Researcher Identifier (e.g. ORCID ID): | ORCID: 0000-0003-3505-3754 |
| Nearest person month worked: | 5 person months |
| Contribution to Project: | Dr. Newcomb facilitates the day-to-day operations of all aspects of the research, interfacing with the PASS Study to ensure high quality data and specimens. She works closely with Dr. Lin and all investigators and collaborators in the execution of the project. Dr. Newcomb is responsible for specimen selection, management of the acquisition and distribution of specimens from the biorepository, as well as overseeing regulatory requirements and supervising study staff. |
| Funding Support: | N/A |

| | |
|---|---|
| Name: | Stephanie Page-Lester |
| Project Role: | QA Coordinator |
| Researcher Identifier (e.g. ORCID ID): | N/A |
| Nearest person month worked: | 1 person month |
| Contribution to Project: | Ms. Page-Lester works closely with Dr. Newcomb, and is responsible for operational activities assuring data quality, protocol adherence, and timeliness of activities. She interfaces between the pathologists and the programmer (Mr. Westcott) in resolving data entry issues related to pathology central review. |
| Funding Support: | N/A |

| | |
|---|---|
| Name: | Deanna Stelling |
| Project Role: | Common Data Element Specialist |
| Researcher Identifier (e.g. ORCID ID): | N/A |
| Nearest person month worked: | 1 person month |
| Contribution to Project: | Ms. Stelling works closely with the programmers and project director to create and maintain the common data elements required for the pathology central review system. She ensures that the data collected by the pathologists are accurately recorded and systematically coded. |
| Funding Support: | N/A |

| | |
|---|---|
| Name: | Maria Tretiakova, MD, PhD |
| Project Role: | Co-investigator, Pathologist |
| Researcher Identifier (e.g. ORCID ID): | ORCID: 0000-0002-0819-9638 |
| Nearest person month worked: | 2 person months |
| Contribution to Project: | Dr. Tretiakova is responsible for reviewing slides of prostate needle biopsies and characterizing the pathologic parameters such as Gleason score and amount of cancer. She is also working with co-investigators at FHCRC and Cleveland Clinic on study design, data analysis, and interpretation. |
| Funding Support: | N/A |

| | |
|---|---|
| Name: | Lawrence True, MD |
| Project Role: | Pathologist |
| Researcher Identifier (e.g. ORCID ID): | ORCID: 0000-0002-8621-9569 |
| Nearest person month worked: | 1 person month |
| Contribution to Project: | Dr. True is responsible for reviewing slides of prostate needle biopsies and characterizing the pathologic parameters such as Gleason score and amount of cancer. He is working with co-investigators at FHCRC and Cleveland Clinic on study design, data analysis, and interpretation. |
| Funding Support: | N/A |

| | |
|---|---|
| Name: | Richard Westcott |
| Project Role: | Programmer |
| Researcher Identifier (e.g. ORCID ID): | N/A |
| Nearest person month worked: | 3 person months |
| Contribution to Project: | Mr. Westcott is responsible for customizing and maintaining the PASS study database as well as the central pathology review system. This includes creation of custom slide views, annotation forms, and reports to facilitate the pathology review workflow and collect and monitor the pathology review data. Mr. Westcott also prepares reports for investigators and the PASS team. |
| Funding Support: | N/A |

| | |
|---|---|
| Name: | Yingye Zheng, PhD |
| Project Role: | Co-investigator, Biostatistician |
| Researcher Identifier (e.g. ORCID ID): | ORCID: 0000-0002-3078-4200 |
| Nearest person month worked: | 2 person months |
| Contribution to Project: | Dr. Zheng is responsible for all statistical aspects of this project, including design and analysis. She consults with investigators on study designs and necessary study design modifications if necessary during the course of the study. She ensures that appropriate data items are collected for valid data analyses and QA/QC to be conducted to ensure high quality of clinical and assay data. She also supervises the SRA in data analyses and interpretation of study data. |
| Funding Support: | N/A |

**Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?**

Yes. Listed below are changes in the other support for senior and key personnel. Please note: none of these changes impacts effort on the project.

**LIN, D.**

New Funding:

U01 CA199338 (Etzioni) 9/1/15 – 8/31/17 0.24 cal months
NIH $22,956
Modeling to Improve Cancer Outcomes Across Diverse Populations (CISNET)

**MCKENNEY, J.**

No changes .

**TRETIAKOVA, M.**

No changes.

**TRUE, L.**

New Funding:

P50 CA97186-11 (Mostaghel) 1/1/16 – 12/31/16 0.12 cal months
NIH/NCI $27,178
Development of an Automated Image Analysis Protocol and Analytic Algorithm to Quantify Relative Levels of AR and AR Variants in Prostate Cancer (Pilot Project)

Funding Ended:

| | | |
|---|---|---|
| W81XWH13-2-0070 (Scher) | 9/30/13 – 9/26/16 | 1.8 cal months |
| DOD | $474,348 | |

Toward the Practice of Precision Medicine: Multicenter Validation of of Biomarker Assays for Clinical Management of Prostate Cancer.

**ZHENG, Y.**

New Funding:

| | | |
|---|---|---|
| RO1 CA195789 (Rubin) | 8/1/14 – 7/31/19 | 0.24 cal months |
| NIH/NCI | $85,800 | |

Precision Medicine Approach to Prostate Cancer Active Surveillance

| | | |
|---|---|---|
| UO1 CA113913 (Sanda) | 4/11/16 – 3/31/21 | 0.25 cal months |
| NIH/NCI | $30,895 | |

Emory, Harvard, and University of Washington Prostate Cancer Biomarker Center

| | | |
|---|---|---|
| UO1 CA185094 (Peters) | 9/1/14 – 8/31/17 | 0.6 cal months |
| NIH/NCI | $375,329 | (effort Yr 3 only) |

**What other organizations were involved as partners?**

**Organization Name:** University of Washington
**Location of Organization**: Seattle, WA
**Partner's contribution to the project:**

**Facilities**: Staff (Drs. Lin, True, Tretiakova) used facilities provided by the University of Washington for pathology review and office space.
**Collaboration**: University of Washington personnel provide expertise in pathology (Drs. Tretiakova and True) and study oversight (Dr. Newcomb).

**Organization Name**: Cleveland Clinic
**Location of Organization**: Cleveland, OH
**Partner's contribution to the project:**

**Facilities:** Dr. McKenney uses facilities provided by the Cleveland Clinic for central pathology review.
**Collaboration:** Dr. McKenney provides expertise for central pathology review.

**Organization Name:** Genomic Health, Inc.
**Location of Organization**: Redwood City, CA
**Partner's contribution to the project:**

**Collaboration**: Genomic Health, Inc. has agreed to run Prostate Oncotype Dx assays free of charge and discussed design of project.

**Organization Name:** OPKO Diagnostics
**Location of Organization**: Miami, FL
**Partner's contribution to the project:**
**Collaboration**: OPKO Diagnostics has run the blood kallikrein assays free of charge and discussed design of project.

**Organization Name:** Hologic GenProbe
**Location of Organization**: San Diego, CA
**Partner's contribution to the project:**
**Collaboration**: Hologic GenProbe has run the PCA3 and TMPRSS2:ERG urine marker assays free of charge.

## 8. SPECIAL REPORTING REQUIREMENTS

**COLLABORATIVE AWARDS**:
For this project, Dr. Daniel Lin is the initiating PI and Dr. Jesse McKenney is the partnering PI. Drs. Lin and McKenney are independently submitting a duplicate annual project report, with tasks clearly marked with the responsible PI and research site as requested.

**QUAD CHARTS**: Not applicable.

## 9. APPENDICES

Appendix 1. 4KScore Manuscript Under Review, Pages 23 – 49.

Daniel W. Lin, Lisa F. Newcomb, Marshall D. Brown, Daniel D. Sjoberg, Yan Dong, James D. Brooks, Peter R. Carroll, Matthew Cooperberg, Atreya Dash, William J. Ellis, Michael Fabrizio, Martin E. Gleave, Todd M. Morgan, Peter S. Nelson, Ian M. Thompson, Andrew Wagner, and Yingye Zheng for the Canary Prostate Active Surveillance Study Investigators. “Evaluating the four kallikrein panel of the 4Kscore for prediction of high-grade prostate cancer in 2 men in the Canary Prostate Active Surveillance Study (PASS).” European Urology. Under review.

# Evaluating the four kallikrein panel of the 4Kscore for prediction of high-grade prostate cancer in men in the Canary Prostate Active Surveillance Study (PASS)

Daniel W. Lin*, Lisa F. Newcomb, Marshall D. Brown, Daniel D. Sjoberg, Yan Dong, James D. Brooks, Peter R. Carroll, Matthew Cooperberg, Atreya Dash, William J. Ellis, Michael Fabrizio, Martin E. Gleave, Todd M. Morgan, Peter S. Nelson, Ian M. Thompson, Andrew Wagner, and Yingye Zheng for the Canary Prostate Active Surveillance Study Investigators.

Author Affiliations: Fred Hutchinson Cancer Research Center (Lin, Newcomb, Brown, Nelson, Zheng); University of Washington (Lin, Newcomb, Ellis); Memorial Sloan-Kettering Cancer Center (Sjoberg), OPKO (Dang), Stanford University (Brooks); University of British Columbia (Gleave); University of California at San Francisco (Carroll, Cooperberg); University of Texas Health Sciences Center at San Antonio (Thompson), Veterans Affairs Puget Sound Health Care System (Lin, Dash); Beth Israel Deaconess Medical Center (Wagner); Eastern Virginia Medical School (Fabrizio); and the University of Michigan (Morgan)

Runninghead: Four kallikrein panel in active surveillance

Keywords: prostatic neoplasms, prospective studies, active surveillance, biomarker, kallikrein
Word count (abstract): 297
Word count (text only): 2,493

Funding sources: Department of Defense (PC130355), Canary Foundation, Institute for Prostate Cancer Research

Correspondence:

Daniel W. Lin, MD
Department of Urology
University of Washington
1959 NE Pacific Street, Box 356510
Seattle, WA 98195
dlin@uw.edu
phone: 206-221-0797
fax: 206-543-3272

## ABSTRACT

**Background:** Diagnosis of Gleason 6 prostate cancer can leave uncertainty about the presence of undetected aggressive disease.

**Objective:** Evaluate utility of a four kallikrein (4K) panel to predict presence of high-grade cancer in men on active surveillance.

**Design, Setting, and Participants:** Plasma collected prior to the first and subsequent surveillance biopsies was assessed in 718 men prospectively enrolled in the multi-institutional Canary PASS. Biopsy data were split 2:1 into training and test sets. Statistical models were developed including clinical information and either the 4Kpanel or serum PSA.

**Outcome Measurements and Statistical Analysis:** The endpoint was reclassification to Gleason ≥7. We used Receiver Operating Characteristic (ROC) curve analyses and area under the curve (AUC) to assess discriminatory capacity and decision curve analysis (DCA) to report clinical net benefit.

**Results and Limitations:** Significant predictors for reclassification were 4Kpanel (OR=1.54 [1.31,1.81]) or PSA (OR=2.11 [1.53,2.91]), ≥20% cores positive (OR=2.10 [1.33,3.32]), ≥2 prior negative biopsies (OR=0.19 [0.04,0.85]), prostate volume (OR=0.47 [0.31,0.70]), BMI (OR=1.09 [1.04,1.14]). ROC curve analysis comparing 4K and base models indicated that the 4Kpanel improved accuracy for predicting reclassification (AUC 0.78 versus 0.74) in the first surveillance biopsy. Both models performed comparably for prediction of reclassification in subsequent biopsies (AUC=0.75 versus 0.76). In DCA, both models showed higher net benefit compared to biopsy-all and biopsy-none strategies. Limitations include that this study was performed in a single cohort with small numbers; results should be validated in another cohort before clinical use.

**Conclusions:** The 4Kpanel provided incremental value over routine clinical information in predicting high-grade cancer in the first biopsy after diagnosis. The 4Kpanel did not add predictive value to the base model at subsequent surveillance biopsies.

**Patient Summary:** Active surveillance is a management strategy for many low-grade prostate cancers. Repeat biopsies monitor for previously undetected high-grade cancer. We show that a model with clinical variables, including a panel of four kallikreins, indicates presence of high-grade cancer before a biopsy is performed.

## INTRODUCTION

Active surveillance is a management strategy for men with low-grade, localized prostate cancer that allows men to delay or be spared the potential morbidities of treatment. Cancers that appear low-risk at diagnosis are monitored, typically with serial PSA measurements, clinical exams, and repeat prostate biopsies. Intervention is recommended for evidence of a more aggressive tumor, usually based on changes in biopsy characteristics.

However, fear of occult high-grade cancer in part due to the known undersampling of systematic prostate biopsies has tempered widespread adoption of active surveillance. Even with emerging MRI-based biopsy protocols, there remains uncertainty surrounding the presence of more aggressive disease in a background of apparently low-risk cancer. Additionally, the optimal surveillance schedule and triggers for intervention have not been established, resulting in substantial variation in the practice of active surveillance. Prostate biopsy can be painful, anxiety-provoking, expensive, and potentially-morbid, thus, avoiding unnecessary surveillance biopsies is attractive. Methods to reduce the number of biopsies in active surveillance regimens, while maximizing the identification of high-grade cancers that may benefit from treatment, would have substantial clinical utility.

A promising approach to determine active surveillance candidacy and surveillance regimens, e.g. more intensive versus less intensive biopsy schedules, involves the addition of biomarker panels to prediction models based on known clinical and demographic variables.[1] In men suspected of having prostate cancer, a panel of four kallikreins (total PSA (tPSA), free PSA (fPSA), intact PSA (iPSA), and human kallikrein 2 (hK2)) combined with age using a mathematical algorithm has been shown to improve the prediction of high-grade cancers compared to the PCPT risk calculator or models using tPSA alone.[2,3] Here, we explore utility of prediction models incorporating the pre-defined 4 kallikrein panel algorithm (4Kpanel) to predict the presence of occult high-grade disease in men already diagnosed with Gleason 6 cancer and on active surveillance. We use plasma specimens and data from the prospective, multi-institutional Canary Prostate Active Surveillance Study (PASS).

## MATERIAL AND METHODS

### *Study Cohort*

This study included men from the Canary Prostate Active Surveillance Study (PASS), a multicenter, prospective study enrolling men on active surveillance.[4] Participants in PASS consented to specimen collection as part of the PASS protocol (clinicaltrials.gov NCT00756665), approved by institutional review boards at participating sites. The PASS protocol includes monitoring at clinic visits every 6 months with the first ≥10 core surveillance prostate needle biopsy at 6-12 months, the second 24 months from cancer diagnosis, and subsequent biopsies every 2 years. Specimens, including EDTA plasma, were collected at study entry and every 6-month clinic visit and stored at -70°C until use.

In February 2015, 1170 participants were enrolled in PASS at nine sites throughout North America. Of these, 956 participants had an on-study biopsy, 877 with Gleason 3+3 disease at study entry, 771 had not used 5-alpha reductase inhibitors, and 753 had available EDTA plasma collected prior to biopsy. Participants with missing prostate volume or ratio of positive to total biopsy cores were excluded from modeling (N=35); the remaining 718 men, who had 1,111 biopsies, were included in this study.

### *Laboratory Methods*

Blood was collected in $K_2$EDTA vacutainers, inverted, centrifuged at 1600g-force and frozen at -70°C within 4 hours of collection. Frozen plasma was stored until shipment on dry ice to OPKO Labs (Nashville, TN) for analysis. The analysis lab was blinded to all specimen and clinical information. Specimens were thawed immediately prior to analysis. tPSA, fPSA, iPSA and hK2 were measured.[2]

### *Study Design and Analyses*

The objective of the analyses was to determine whether a model using clinical predictors and kallikrein data collected after diagnosis of Gleason 6 cancer, but prior to surveillance biopsy, can predict high-grade cancer in the surveillance biopsy. Sequential surveillance biopsies were considered as two groups: A) the

initial biopsy after cancer diagnosis (sometimes called confirmatory biopsy), and B) all subsequent surveillance biopsies. Biopsy data were split 2:1 into training and test sets matched by outcome.

The primary outcome was reclassification from Gleason score 6 to Gleason score ≥7. A value for the 4Kpanel was calculated with tPSA, fPSA, iPSA, hk2 and age using locked down coefficients developed before the study was conducted.[3] This combination of the four kallikreins was the same as in the commercial 4Kscore. However the commercial 4Kscore is a model containing the 4Kpanel and clinical data available prior to cancer diagnosis and is calibrated for a patient prior to diagnosis. Because we evaluated the kallikreins in a cohort already diagnosed with cancer, we developed a new model that included the 4Kpanel and clinical information available after a diagnosis of cancer and is calibrated to an active surveillance population. Additional clinical predictors considered in modeling included age, body mass index (BMI), race (African American or other), digital rectal examination (DRE) results, number of previous biopsies after diagnosis, number of negative biopsies after diagnosis, cores ratio (ratio of biopsy cores containing cancer to total cores) from previous biopsy, maximum cores ratio from all previous biopsies, months since diagnosis, prostate volume (prostate size measured closest to time of sampling and imputed within 2-years). Either the 4Kpanel (logit scale) or clinical serum PSA (logarithm transformed) was used in models. Prediction models were built using data in the training set and clinical performance was assessed with the testing set. We follow the principles set forth by the FDA Critical Path Initiative, using an established biomarker with analytic validity for the intent of clinical validation in the intended use population.[7] Furthermore, we follow reporting recommendations for tumor marker prognostic studies (REMARK)[8] and the Tumor Marker Utility Grading System (TMUGS)[9] in reporting the clinical utility of the biomarker panel.

Model building: Data from initial and subsequent biopsy groups were combined for model development. Interaction terms of biopsy group indication (initial vs. subsequent surveillance biopsy) with other variables were evaluated to investigate whether effects may differ for an initial biopsy and subsequent biopsy. Logistic regression was used to fit the models, with robust variance to account for the correlation among multiple biopsies on the same patient. Forward stepwise model selection procedures were

implemented. Variable selection criteria included p-values <0.15, area under the Receiver Operating Characteristic(ROC) Curve(AUC) ≥0.005, or a quasilikelihood under the independence model criterion(QIC) with threshold =0.[5] Final models were compared to identify variables that were robust to selection procedures. We first identified a full model including clinical predictors and 4Kpanel, and then a base model with serum PSA substituted for the 4Kpanel. In some clinics, prostate volume may not be reliably available, thus, models without prostate volume were fit sequentially.

Model validation: Calibration plots were used to gauge the goodness of fit of each model. We used ROC analyses and AUC to assess the discriminatory capacity of a model for separating patients with and without reclassification. Decision curve analysis (DCA) was used to report the clinical net benefit of each model compared to biopsy-all and biopsy-none strategies.[6] The potential clinical impact was illustrated by plotting number of cancers missed versus the number of biopsies avoided per 1,000 individuals. To illustrate the clinical consequence of each model, we report the number of biopsies that could be avoided and the number of Gleason ≥7 cancers that might be missed if a risk-based threshold is applied as a criterion for biopsy. All evaluations were done on the initial biopsy and subsequent biopsy groups separately and combined. Confidence intervals and significant tests were calculated using the Bootstrap resampling procedure to account for within-subject correlations.

All Analyses were conducted using R version 3.1.1 (http://www.r-project.org/).

## RESULTS

Of the 718 men in this study, there were 478 participants in the initial biopsy group for whom kallikreins were assayed: 319 in the training set [60 (18.8%) with Gleason ≥7] and 159 in the test set [34 (21.4%) with Gleason ≥7] (Table 1). In bivariate analyses, prostate volume, ratio of positive to total cores, and the 4Kpanel were significantly associated with grade reclassification. There were 444 participants (of which 204 were also in the initial biopsy group) with 633 subsequent surveillance biopsies, 422 in the training set [70 (17%) with Gleason ≥7] and 211 in the test set [31 (15%) with Gleason ≥7] (Table 2, Supplementary Table 1 respectively). Biopsies in this group ranged from the second to eighth after

diagnosis and most patients had Gleason score 6 or no cancer at their surveillance biopsies, varying slightly across biopsy number.

In the full clinical model (Table 3) including the 4Kpanel, significant predictors for reclassification were BMI (odds ratio(OR) = 1.09 [95% CI: 1.04,1.14]), more than 20% of cores positive in the prior biopsy (OR = 2.10 [95% CI: 1.33,3.32]), a history of 2 or more biopsies negative for cancer (OR = 0.19 [95% CI: 0.04,0.85]), prostate volume (per fold increase, OR = 0.47 [95% CI: 0.31,0.70]), and 4Kpanel (OR = 1.5 [95% CI: 1.31,1.81]). In the clinical model with serum PSA replacing the 4Kpanel, PSA was significantly associated with reclassification (per fold increase, OR = 2.11 [95% CI: 1.53,2.91]), and age was not. In models that did not include prostate volume, the effects were similar for covariates left in the model (Supplementary Table 2). Model calibration in the test set showed predicted probabilities of reclassification closely matching the empirical rates (Supplementary Figure 1).

ROC curve analysis (Table 4, Supplementary Figure 2) comparing the full model with 4Kpanel and the full clinical model with serum PSA indicated that the 4Kpanel significantly improved the accuracy for predicting reclassification (AUC 0.78 versus 0.74) in the initial surveillance biopsy, with a significant incremental value in AUC of 0.04 (95% CI: 0.003,0.09). In a model without prostate volume, the incremental value in AUC is 0.07 (95% CI: 0.02,0.11). The 4Kpanel did not improve prediction of reclassification in subsequent biopsies relative to PSA (AUC 0.75 with 4Kpanel versus 0.76 with PSA).

Similar findings were observed in DCA. Compared to a clinical model with serum PSA, the model with 4Kpanel showed a higher net benefit for the initial surveillance biopsy, but there was no benefit for subsequent biopsies. All models showed substantial gain in net benefit compared with the biopsy-all and biopsy-none strategies across a range of plausible cost and benefit ratios (Figure 1 and Supplementary Figure 3).

The clinical consequences, or the number of biopsies and the number of high-grade cancers that could be avoided or delayed per 1,000 AS patients, were illustrated based on prediction models with 4Kpanel or PSA (Table 5). For example, using a model with 4K and a clinical rule of only performing an initial surveillance biopsy in patients whose risk of high-grade cancer exceeded 10%, 252 biopsies would be

avoided, 19 of which would contain high-grade cancer as defined by any pattern 4 disease, and 0 biopsies with primary Gleason 4. Comparing the two models at the same numbers of biopsies avoided (Supplementary Figure 4) shows that the 4K model appears to miss fewer higher grade cancers while avoiding the same number of initial biopsies.

## DISCUSSION

In this study from a prospectively enrolled multi-institutional cohort of men on active surveillance, we show that a panel of four kallikrein markers, when added to a model that includes clinical information, can significantly improve the prediction of the outcome in the first surveillance biopsy. Both models performed comparably for prediction of reclassification in subsequent biopsies. Importantly, in DCA, both models showed higher net benefit compared to biopsy-all and biopsy-none strategies. Lastly, we showed how the 4K panel added to currently available clinical metrics and how the results impact clinical management.

There is a growing body of evidence that true Gleason 6 prostate cancer is indolent and will not cause harm if left untreated.[10-12] This knowledge is balanced by the known undersampling in prostate needle biopsies, and while some have promoted that select Gleason 3+4 cancers may undergo surveillance, level I clinical trial data and treatment guidelines generally recommend treatment of higher grade cancers, including Gleason 3+4 disease.[13,14] Our efforts focus on developing tools for use after the diagnosis of Gleason 6 prostate cancer to provide a higher degree of certainty that no occult high-grade cancer was missed at diagnosis. More accurate tools would not only support the practice of active surveillance, but could also promote less intensive monitoring regimens.

A panel of four kallikreins, when combined in a mathematical algorithm, has been shown to improve the prediction of newly diagnosed high-grade (Gleason ≥7) cancer.[3] This panel of markers has also been shown to improve long-term prediction of metastatic disease in men with PSA ≥2 in a Swedish cohort.[15] In this study, we asked whether the same panel of markers[3] improved the prediction of high-grade disease in surveillance biopsies of men already diagnosed with Gleason 6 cancer. We found that when the kallikreins were assessed prior to the initial surveillance biopsy (sometimes called the confirmatory biopsy), the

4Kpanel provides incremental benefit for prediction of high-grade cancer (Gleason ≥7) over the clinical factors that are available at diagnosis. Specifically, depending on the choice from the various cutpoints that are based on the risk of high-grade disease, a substantial number of biopsies could be avoided, while minimizing the number of missed high-grade cancers, few of which had primary pattern 4. The 4Kpanel was not of value over PSA for the prediction of reclassification in subsequent biopsies after the first surveillance biopsy. We found that the impact of other biopsy information, primarily volume of core involvement in previous biopsies and the number of previous negative biopsies, carries such statistical weight in modeling that the impact of the 4Kpanel is minimized. For example, if a patient had low volume disease at the initial surveillance biopsy, or if the patient had subsequent negative biopsies after the initial diagnosis, then these factors were highly protective against biopsy reclassification at subsequent biopsy. It should be noted that our analysis of these subsequent biopsies used the 4Kpanel from the plasma sample that was closest to the subsequent biopsy, not necessarily the plasma sample from study entry which could be months or years earlier than the subsequent biopsy.

We include serum PSA and prostate volume separately in our models, instead of calculating PSA density, as we find better model fit when the variables enter the model independently. While TRUS prostate volume measurements may have imprecision,[16] statistical models that included prostate volume appeared to provide slightly improved predictive performance (AUC for all groups 0.77 with volume versus 0.75 without volume). Furthermore, prostate volume is a strong predictor of finding higher grade cancers, with larger prostates being protective, a previously reported finding.[17]

This study has limitations that merit mention. First, the model was developed and tested in the same cohort and with relatively limited numbers that resulted in wide confidence intervals and minor differences between the training and test sets. The results should clearly be validated in other cohort before clinical application. However, we expect that our results will be similar to those found in a community setting as PASS is multicenter center study that represents a broad spectrum of men utilizing active surveillance. Similarly, as PASS is primarily a Caucasian cohort, the findings of this study may not be generalizable to African American patients. Another limitation is that the serum PSA measurements used were obtained as

part of standard clinical care, and the local site assays may differ from the one used with the 4Kpanel. Thus, the comparative modeling using PSA versus 4Kpanel may have slightly different tPSA values with caution suggested for comparisons between the models. Lastly, as the use of imaging, such as multiparametric MRI (mpMRI), is increasing, we do not have MRI data on most of our participants and recognize the potential value of future studies incorporating results from mpMRI and biomarkers in active surveillance.

## CONCLUSION

The 4Kpanel was significantly associated with reclassification at the first surveillance biopsy, providing incremental value over routine clinical information, and the 4K model performed significantly better than the base model in this group. The 4Kpanel did not add predictive value to a PSA clinical model for biopsy decision-making with men at subsequent surveillance biopsies. This work aims to provide clinical validation of a biomarker that will help determine those men who have or will develop aggressive prostate cancer, allowing for the accurate determination of those men who may avoid or delay the burden of immediate treatment safely, while concurrently identifying men who may optimally benefit from early treatment.

## References


1. Ankerst DP, Xia J, Thompson IM, Jr., et al. Precision Medicine in Active Surveillance for Prostate Cancer: Development of the Canary-Early Detection Research Network Active Surveillance Biopsy Risk Calculator. *Eur Urol.* 2015;68(6):1083-1088.
2. Parekh DJ, Punnen S, Sjoberg DD, et al. A multi-institutional prospective trial in the USA confirms that the 4Kscore accurately identifies men with high-grade prostate cancer. *Eur Urol.* 2015;68(3):464-470.
3. Bryant RJ, Sjoberg DD, Vickers AJ, et al. Predicting high-grade cancer at ten-core prostate biopsy using four kallikrein markers measured in blood in the ProtecT study. *J Natl Cancer Inst.* 2015;107(7).
4. Newcomb LF, Thompson IM, Jr., Boyer HD, et al. Outcomes of Active Surveillance for Clinically Localized Prostate Cancer in the Prospective, Multi-Institutional Canary PASS Cohort. *J Urol.* 2016;195(2):313-320.
5. Pan W. Akaike's information criterion in generalized estimating equations. *Biometrics.* 2001;57(1):120-125.
6. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* 2006;26(6):565-574.
7. Report USDoHaHSFaDA. *Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products.* 2004.
8. McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst.* 2005;97(16):1180-1184.
9. Hayes DF, Bast RC, Desch CE, et al. Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers. *J Natl Cancer Inst.* 1996;88(20):1456-1466.
10. Ross HM, Kryvenko ON, Cowan JE, Simko JP, Wheeler TM, Epstein JI. Do adenocarcinomas of the prostate with Gleason score (GS) </=6 have the potential to metastasize to lymph nodes? *Am J Surg Pathol.* 2012;36(9):1346-1352.
11. Eggener SE, Badani K, Barocas DA, et al. Gleason 6 Prostate Cancer: Translating Biology into Population Health. *J Urol.* 2015;194(3):626-634.
12. Carter HB, Partin AW, Walsh PC, et al. Gleason score 6 adenocarcinoma: should it be labeled as cancer? *J Clin Oncol.* 2012;30(35):4294-4296.
13. Wilt TJ, Brawer MK, Jones KM, et al. Radical prostatectomy versus observation for localized prostate cancer. *N Engl J Med.* 2012;367(3):203-213.
14. Mohler JL, Kantoff PW, Armstrong AJ, et al. Prostate cancer, version 1.2014. *J Natl Compr Canc Netw.* 2013;11(12):1471-1479.
15. Stattin P, Vickers AJ, Sjoberg DD, et al. Improving the Specificity of Screening for Lethal Prostate Cancer Using Prostate-specific Antigen and a Panel of Kallikrein Markers: A Nested Case-Control Study. *Eur Urol.* 2015;68(2):207-213.
16. Rodriguez E, Jr., Skarecky D, Narula N, Ahlering TE. Prostate volume estimation using the ellipsoid formula consistently underestimates actual gland size. *J Urol.* 2008;179(2):501-503.
17. Roobol MJ, van Vugt HA, Loeb S, et al. Prediction of prostate cancer risk: the role of prostate volume and digital rectal examination in the ERSPC risk calculators. *Eur Urol.* 2012;61(3):577-583.

**Table 1.**

| Characteristics | Training Set | | | Test Set | | |
|---|---|---|---|---|---|---|
| | Gleason < 7 | Gleason ≥ 7 | p-value | Gleason < 7 | Gleason ≥ 7 | p-value |
| **Sample size** | 259 | 60 | | 125 | 34 | |
| **Age at diagnosis** | 63 (58–67) | 64 (60–68) | 0.109 | 64 (58–68) | 64 (57–67) | 0.876 |
| **BMI** | 27 (25–30) | 28 (25–33) | 0.116 | 27 (25–29) | 28 (26–31) | 0.305 |
| **Race** | | | | | | |
| Non-African American | 248 (96) | 56 (93) | | 121 (97) | 29 (85) | |
| African American | 11 (4) | 4 (7) | 0.646 | 4 (3) | 5 (15) | 0.522 |
| **Months from diagnosis** | 12.0 (8.4–14.1) | 12.7 (8.6–14.8) | 0.237 | 12.2 (8.8–14.0) | 12.6 (10.3–17.6) | 0.189 |
| **DRE** | | | | | | |
| Normal | 238 (92) | 55 (92) | | 118 (94) | 30 (88) | |
| Abnormal | 21 (8) | 5 (8) | 0.971 | 7 (6) | 4 (12) | 0.031 |
| **Prostate volume (cc)** | 41.0 (30.0–56.5) | 35.5 (25.0–50.0) | 0.041 | 40.0 (30.0–51.0) | 30.0 (24.0–42.8) | 0.006 |
| **Ratio of positive to total cores** | 0.08 (0.08–0.17) | 0.17 (0.08–0.20) | <0.001 | 0.08 (0.08–0.17) | 0.17 (0.17–0.25) | <0.001 |
| **Clinical serum PSA (ng/ml)** | 4.60 (2.91–6.40) | 4.81 (4.35–6.42) | 0.108 | 4.56 (3.11–6.24) | 5.65 (4.58–7.88) | 0.024 |
| **4Kpanel (logit)** | 0.21 (0.08-0.29) | 0.32 (0.16-0.44) | <0.001 | 0.20 (0.07-0.28) | 0.36 (0.18-0.53) | <0.001 |

**Table 2.**

| | Initial Biopsy | Subsequent Surveillance Biopsies | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Characteristics** | **First** | **Second** | **Third** | **Fourth** | **Fifth** | **Sixth** | **Seventh** | **Eighth** |
| **Number of biopsies** | 319 | 246 | 108 | 34 | 20 | 10 | 3 | 1 |
| **Core ratio of previous biopsy[a]** | | | | | | | | |
| Median (IQR) | 0.08 (0.08) | 0.07 (0.17) | 0.08 (0.17) | 0.06 (0.12) | 0.06 (0.12) | 0 (0.07) | 0.11 (0.06) | 0 (0) |
| Missing N(%) | 0 | 5 (2) | 5 (5) | 0 | 0 | 0 | 0 | 0 |
| **Max core ratio[b]** | | | | | | | | |
| Median (IQR) | 0.08 (0.08) | 0.11 (0.08) | 0.13 (0.15) | 0.17 (0.13) | 0.10 (0.17) | 0.14 (0.15) | 0.17 (0.08) | 0.17 (0.00) |
| **Number of negative biopsies[c] N (%)** | | | | | | | | |
| 0 | 319 (100) | 145 (59) | 44 (41) | 10 (29) | 4 (20) | 1 (10) | 1 (33) | 0 |
| 1 | 0 | 101 (41) | 38 (35) | 13 (38) | 6 (30) | 3 (30) | 2 (67) | 0 |
| 2 | 0 | 0 | 26 (24) | 6 (18) | 3 (15) | 1 (10) | 0 | 1 (100) |
| 3 | 0 | 0 | 0 | 5 (15) | 2 (10) | 3 (30) | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 5 (25) | 2 (20) | 0 | 0 |
| **Prostate volume (cc)** | | | | | | | | |
| Median (IQR) | 41.0 (26.5) | 38.0 (27.0) | 41.0 (27.0) | 48.5 (25.0) | 59.5 (36.5) | 43.5 (27.8) | 41.0 (19.5) | 97.0 (0.0) |
| **Gleason score of biopsy N (%)** | | | | | | | | |
| Negative | 107 (34) | 95 (39) | 38 (35) | 11 (32) | 8 (40) | 6 (60) | 2 (67) | 0 |
| 6 | 152 (48) | 108 (44) | 48 (45) | 21 (62) | 10 (50) | 3 (30) | 1 (33) | 1 (100) |
| 7 | 58 (18) | 42 (17) | 21 (19) | 2 (6) | 2 (10) | 1 (10) | 0 | 0 |
| 8 | 1 (0) | 1 (0) | 1 (1) | 0 | 0 | 0 | 0 | 0 |
| 9 | 1 (0) | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

[a] Core ratio is defined as the number of biopsy cores containing cancer divided by the total number of biopsy cores in previous biopsy.

[b] Maximum core ratio in all previous biopsies.

[c] Number of surveillance biopsies in which no cancer was found.

**Table 3:**

| Variable | PSA + full clinical model | | | 4K + full clinical model | | |
|---|---|---|---|---|---|---|
| | OR | CI | p-value | OR | CI | p-value |
| Age | 1.03 | (1.00,1.06) | 0.068 | | | |
| BMI | 1.11 | (1.06,1.16) | <0.001 | 1.09 | (1.04,1.14) | <0.001 |
| Cores ratio >0.2 | 2.19 | (1.39,3.44) | 0.001 | 2.10 | (1.33,3.32) | 0.001 |
| Negative biopsies ≥2 | 0.19 | (0.04,0.80) | 0.023 | 0.19 | (0.04,0.85) | 0.029 |
| Log(prostate volume) | 0.31 | (0.20,0.48) | <0.001 | 0.47 | (0.31,0.70) | <0.001 |
| Log(PSA) | 2.11 | (1.53,2.91) | <0.001 | | | |
| 4Kpanel | | | | 1.54 | (1.31,1.81) | <0.001 |

**Table 4.**

| Base Model | 4K + Clinical Model AUC (95% CI) | PSA + Clinical Model AUC (95% CI) | Difference (95% CI) |
|---|---|---|---|
| Full Clinical Model | | | |
| Initial Biopsy | 0.783 (0.691,0.871) | 0.740 (0.652,0.828) | 0.043 (0.003,0.086) |
| Subsequent Biopsy | 0.754 (0.657,0.838) | 0.755 (0.653,0.841) | -0.001 (-0.037,0.041) |
| Clinical Model without prostate volume | | | |
| Initial Biopsy | 0.748 (0.654,0.840) | 0.678 (0.579,0.774) | 0.069 (0.016,0.114) |
| Subsequent Biopsy | 0.738 (0.633,0.825) | 0.718 (0.611,0.810) | 0.02 (-0.023,0.07) |

**Figure 1.**

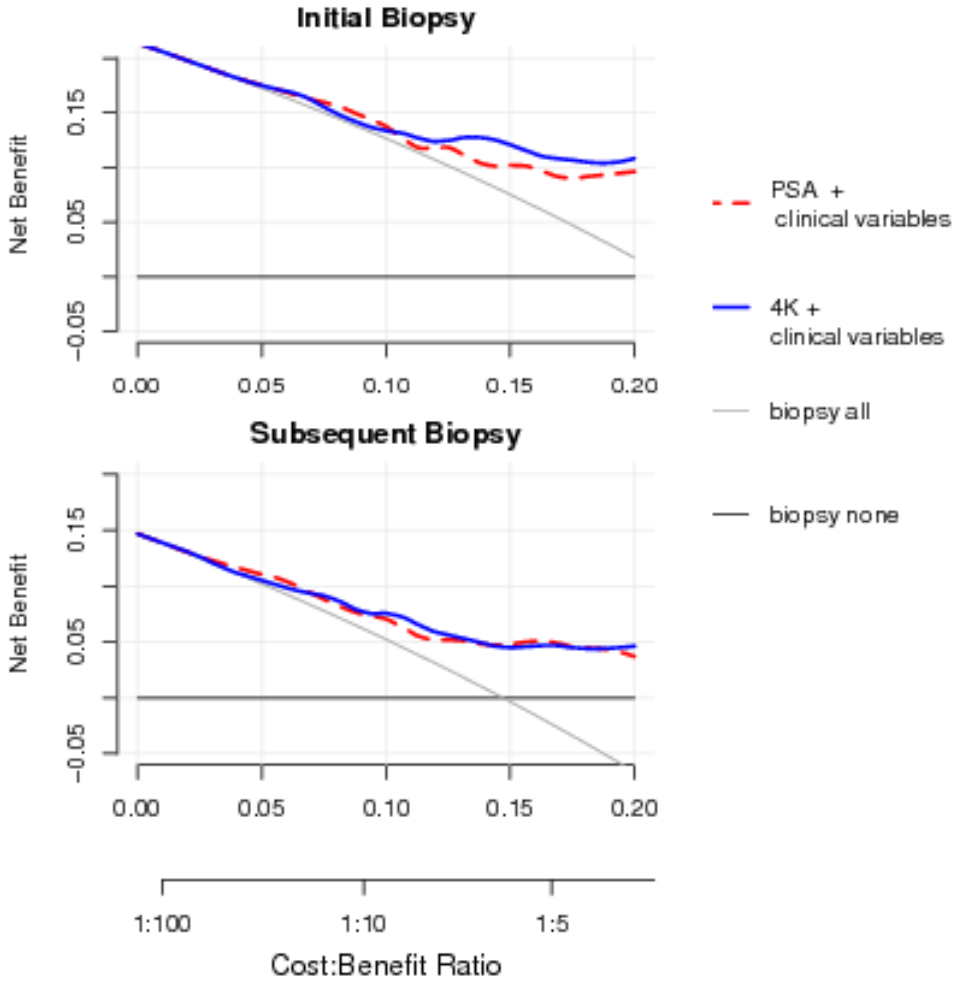

Initial Biopsy
Net Benefit
0.15
0.05
−0.05
0.00
0.05
0.10
0.15
0.20
Subsequent Biopsy
Net Benefit
0.15
0.05
−0.05
0.00
0.05
0.10
0.15
0.20
1:100
1:10
1:5
Cost:Benefit Ratio
PSA + clinical variables
4K + clinical variables
biopsy all
biopsy none

**Table 5.**

| Probability of High-grade Cancer | Biopsies performed | Biopsies Avoided | High-grade Cancers | | Primary Gleason 4 Cancers | |
|---|---|---|---|---|---|---|
| | | | Found | Missed | Found | Missed |
| **Initial Surveillance Biopsy** | | | | | | |
| **Biopsy All** | 1000 | 0 | 214 | 0 | 44 | 0 |
| **Initial Biopsy: Risk by clinical variables + PSA** | | | | | | |
| > 5% | 943 (896, 970) | 57 (30, 104) | 214 (157, 284) | 0 (0, 24) | 44 (21, 88) | 0 (0, 24) |
| > 10% | 761 (689, 821) | 239 (179, 311) | 201 (146, 270) | 13 (3, 45) | 44 (21, 88) | 0 (0, 24) |
| > 15% | 509 (432, 586) | 491 (414, 568) | 164 (114, 229) | 50 (26, 96) | 38 (17, 80) | 6 (1, 35) |
| **Initial Biopsy: Risk by clinical variables + 4K** | | | | | | |
| > 5% | 956 (912, 979) | 44 (21, 88) | 214 (157, 284) | 0 (0, 24) | 44 (21, 88) | 0 (0, 24) |
| > 10% | 748 (676, 809) | 252 (191, 324) | 195 (141, 263) | 19 (6, 54) | 44 (21, 88) | 0 (0, 24) |
| > 15% | 522 (445, 598) | 478 (402, 555) | 182 (130, 250) | 31 (14, 71) | 44 (21, 88) | 0 (0, 24) |
| **Subsequent Surveillance Biopsies** | | | | | | |
| **Biopsy All** | 1000 | 0 | 147 | 0 | 47 | 0 |
| **Risk by clinical variables + PSA** | | | | | | |
| > 5% | 844 (789, 886) | 156 (114, 211) | 147 (105, 201) | 0 (0, 18) | 47 (26, 85) | 0 (0, 18) |
| > 10% | 692 (627, 750) | 308 (250, 373) | 133 (93, 185) | 14 (5, 41) | 43 (23, 79) | 5 (1, 26) |
| > 15% | 445 (380, 513) | 555 (487, 620) | 109 (74, 158) | 38 (19, 73) | 43 (23, 79) | 5 (1, 26) |
| **Risk by clinical variables + 4K** | | | | | | |
| > 5% | 848 (794, 890) | 152 (110, 206) | 142 (101, 196) | 5 (1, 26) | 47 (26, 85) | 0 (0, 18) |
| > 10% | 654 (588, 715) | 346 (285, 412) | 133 (93, 185) | 14 (5, 41) | 47 (26, 85) | 0 (0, 18) |
| > 15% | 408 (344, 475) | 592 (525, 656) | 100 (66, 147) | 47 (26, 85) | 38 (19, 73) | 9 (3, 34) |

**Table and Figure legends**

**Table 1.** Characteristics for 478 participants with kallikreins assayed prior to the initial surveillance biopsy after diagnosis. Medians (with 25-75 percentiles) or counts (percentage of sample size) shown for combined Gleason score <7 versus ≥7 for the training and test cohorts. The value for the 4Kpanel is a predetermined combination of the 4 kallikreins, in the logit scale, that is similar to a component of the commercial 4Kscore.

**Table 2.** Biopsy characteristics at each sequential surveillance biopsy after diagnosis in 558 participants in the training set. Medians with interquartile range (IQR) or numbers with percent (%) are shown.

**Table 3:** Summary of fitted models including clinical variables + serum PSA or 4Kpanel in the training set.

**Table 4.** Results of final regression models for reclassification. AUC (95% CI) of various models for initial surveillance biopsy and subsequent surveillance biopsies. CIs were calculated with bootstrap accounting for correlations among individuals.

**Table 5.** Clinical Consequences showing the number of biopsies that could be avoided for initial surveillance biopsy or subsequent surveillance biopsy. Results are presented per 1000 men. Biopsy numbers and 95% CI are shown.

**Figure 1.** Decision Curve Analysis for full models with serum PSA (dotted red line) or with the 4Kpanel (solid blue line). Strategies for biopsying all men (light grey line) or no men (dark grey line) are also shown. The line with the highest net benefit at any particular threshold probability for biopsy (x-axis) will result in the best clinical results.

**PASS-4K Supplementary Material**

**Supp Table 1.**

| | Initial Biopsy | Subsequent Surveillance Biopsies | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Characteristics of Biopsy History** | **First** | **Second** | **Third** | **Fourth** | **Fifth** | **Sixth** | **Seventh** | **Eighth** |
| **Patient Number** | 159 | 126 | 45 | 25 | 9 | 3 | 2 | 1 |
| **Core ratio of previous biopsy[a]** | | | | | | | | |
| Median (IQR) | 0.10 (0.08) | 0.08 (0.17) | 0 (0.11) | 0.00 (0.12) | 0.06 (0.12) | 0.06 (0.04) | 0.13 (0.13) | 0 (0) |
| Missing N(%) | 0 | 1 (1) | 2 (4) | 0 | 0 | 0 | 0 | 0 |
| **Max core ratio of previous biopsies[b]** | | | | | | | | |
| Median (IQR) | 0.10 (0.08) | 0.13 (0.12) | 0.14 (0.08) | 0.17 (0.10) | 0.20 (0.11) | 0.08 (0.04) | 0.13 (0.13) | 0.08 (0.00) |
| **Number of prior negative biopsies[c] N (%)** | | | | | | | | |
| 0 | 159 (100) | 76 (60) | 14 (31) | 6 (24) | 4 (45) | 0 | 0 | 0 |
| 1 | 0 | 50 (40) | 17 (38) | 8 (32) | 2 (22) | 1 (33) | 0 | 0 |
| 2 | 0 | 0 | 14 (31) | 5 (20) | 3 (33) | 0 | 0 | 1 (100) |
| 3 | 0 | 0 | 0 | 6 (24) | 0 | 1 (33) | 1 (50) | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 1 (33) | 1 (50) | 0 |
| **Prostate volume (cc)** | | | | | | | | |
| Median (IQR) | 39.0 (20.0) | 40.5 (24.0) | 39.0 (27.0) | 41.0 (25.0) | 43.0 (29.0) | 71.0 (72.5) | 50.5 (34.5) | 19.0 (0.0) |
| **Gleason score of biopsy N (%)** | | | | | | | | |
| Negative | 52 (33) | 41 (33) | 17 (38) | 10 (40) | 1 (11) | 0 | 1 (50) | 0 |
| 6 | 73 (46) | 67 (53) | 21 (47) | 11 (44) | 6 (67) | 3 (100) | 1 (50) | 1 (100) |
| 7 | 34 (21) | 17 (13) | 5 (11) | 3 (12) | 2 (22) | 0 | 0 | 0 |
| 8 | 0 | 1 (1) | 2 (4) | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 1 (4) | 0 | 0 | 0 | 0 |

[a] Core ratio is defined as the number of biopsy cores containing cancer divided by the total number of biopsy cores in previous biopsy.

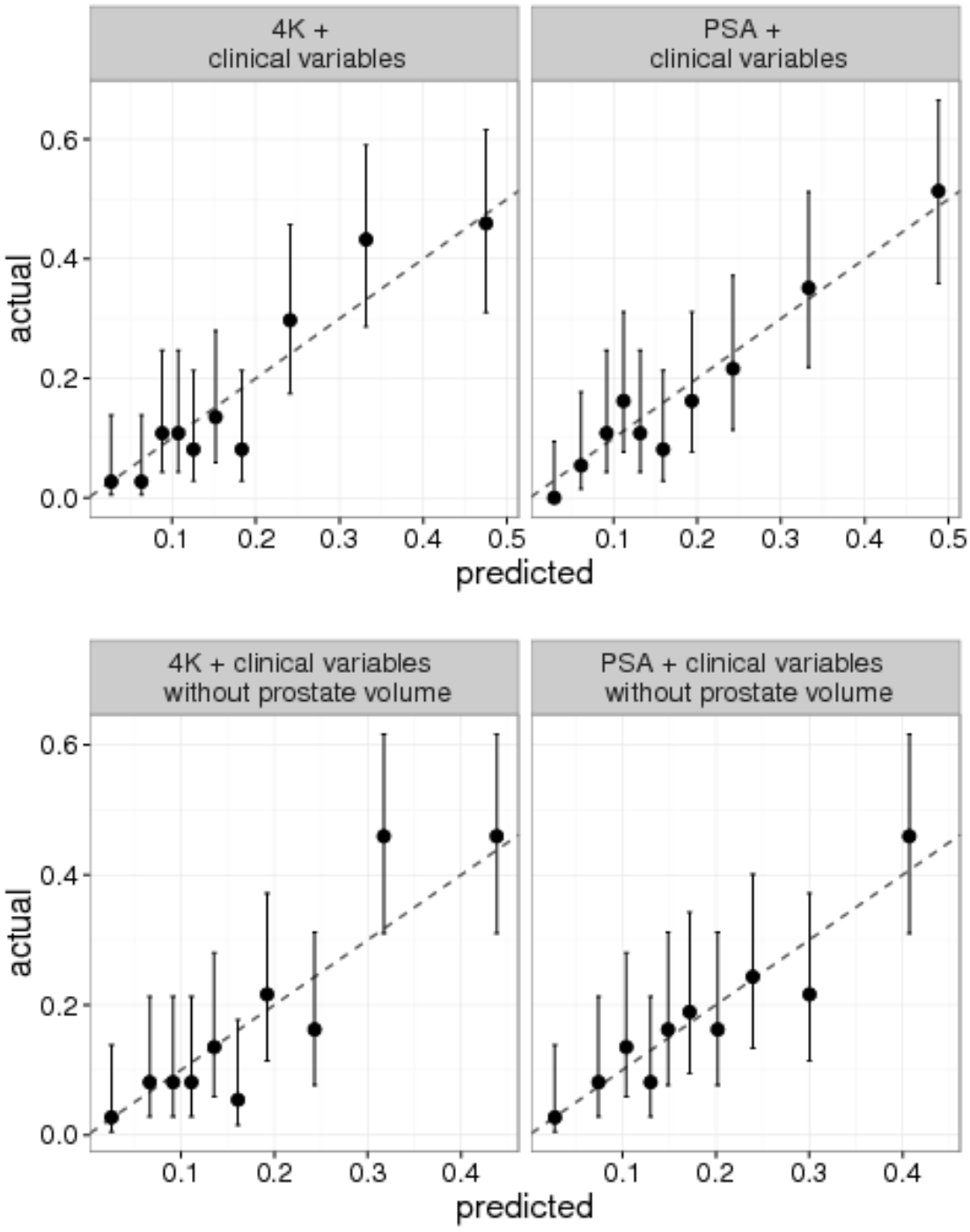
[b] Maximum core ratio in all previous biopsies.

[c] Number of surveillance biopsies in which no cancer was found.

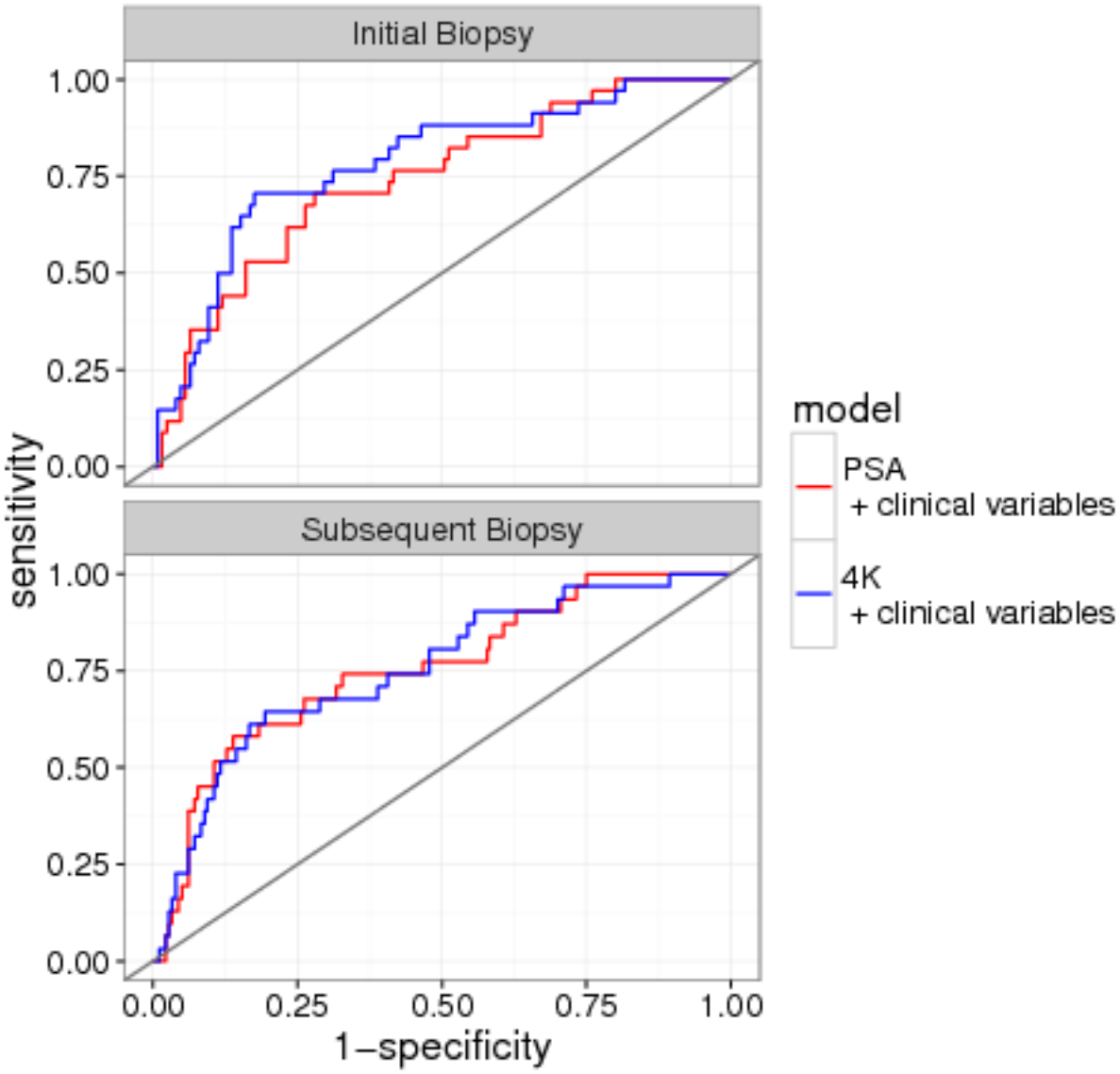
**Supp Table 2:**

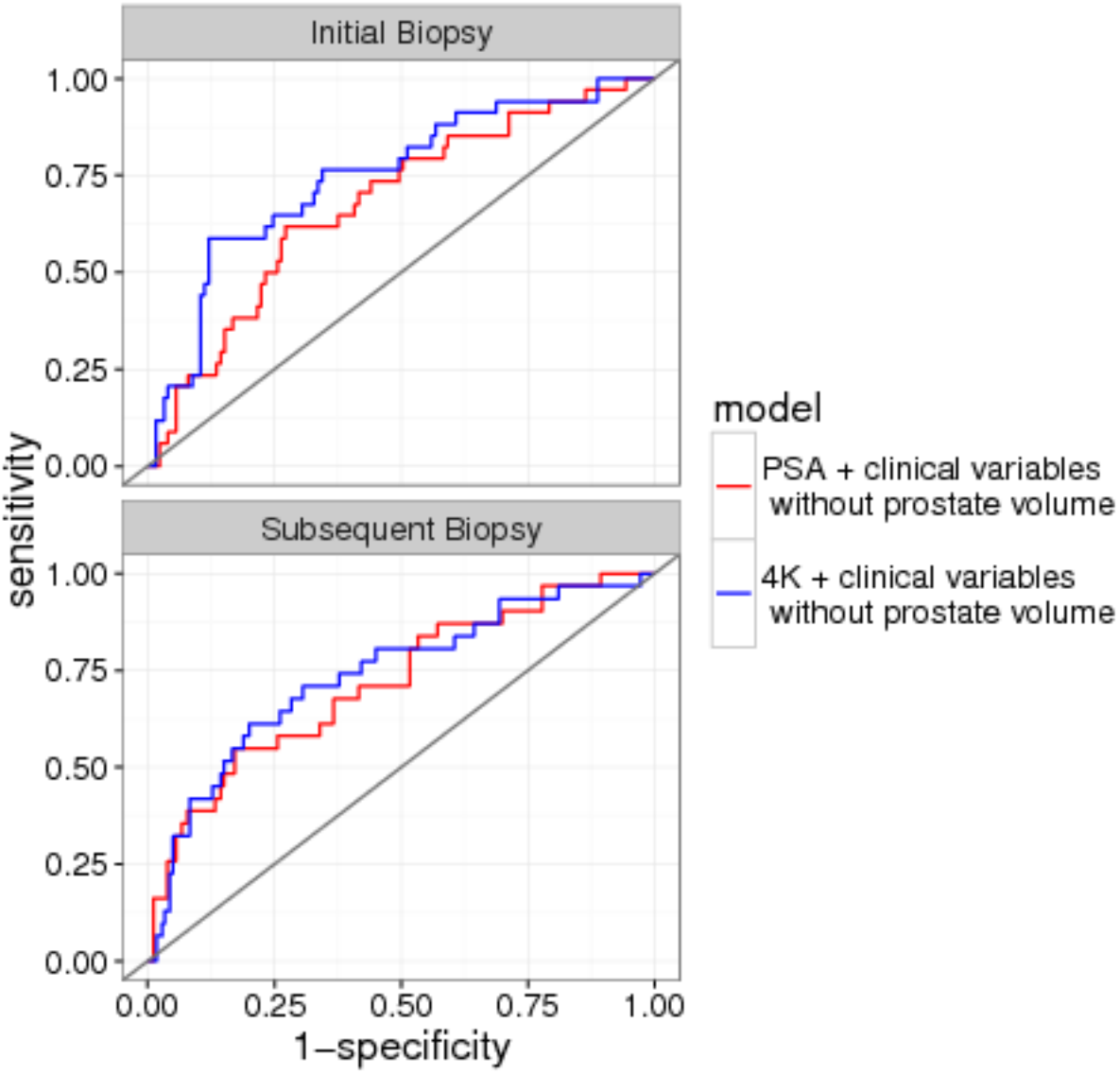
| Variable | PSA + clinical model without prostate volume | | | 4K + clinical model without prostate volume | | |
|---|---|---|---|---|---|---|
| | OR | CI | p-value | OR | CI | p-value |
| Age | 1.02 | (0.99,1.06) | 0.171 | | | |
| BMI | 1.08 | (1.03,1.13) | 0.001 | 1.08 | (1.03,1.13) | 0.001 |
| Cores Ratio >0.2 | 2.58 | (1.66,4.02) | <0.001 | 2.28 | (1.44,3.61) | <0.001 |
| Previous Negative Biopsies ≥2 | 0.15 | (0.03,0.67) | 0.013 | 0.16 | (0.04,0.74) | 0.019 |
| Log(PSA) | 1.65 | (1.21,2.27) | 0.002 | | | |
| 4K panel | | | | 1.54 | (1.31,1.81) | <0.001 |

**Supp Figure 1.**

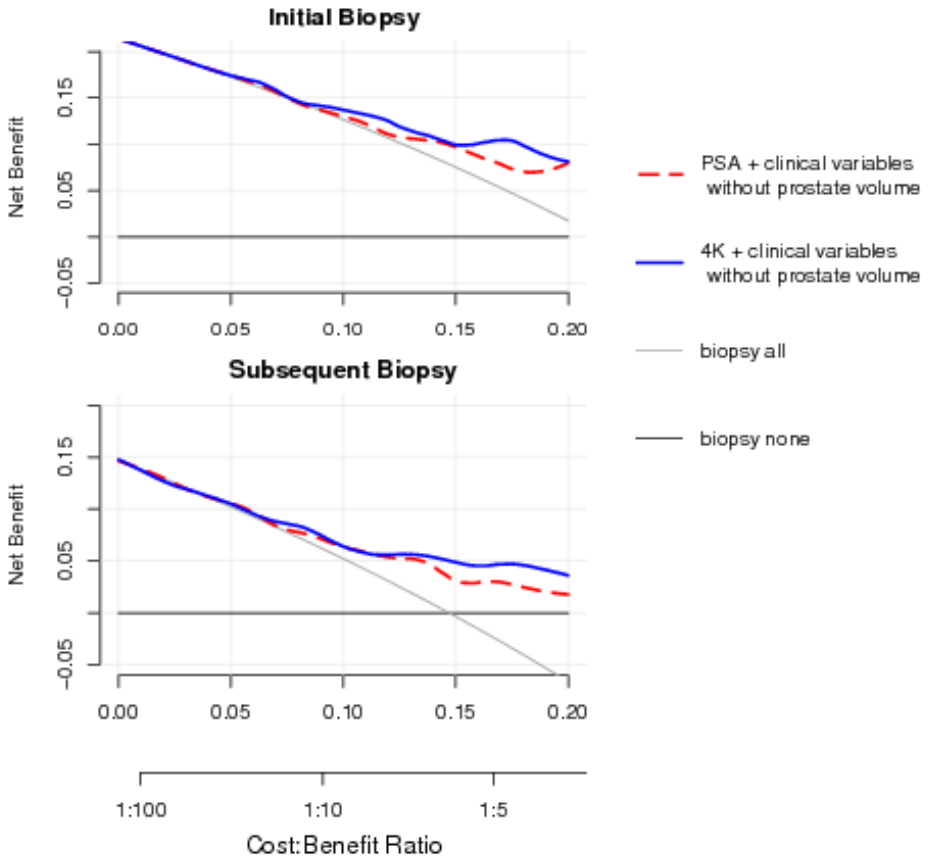


**Supp Figure 2.**

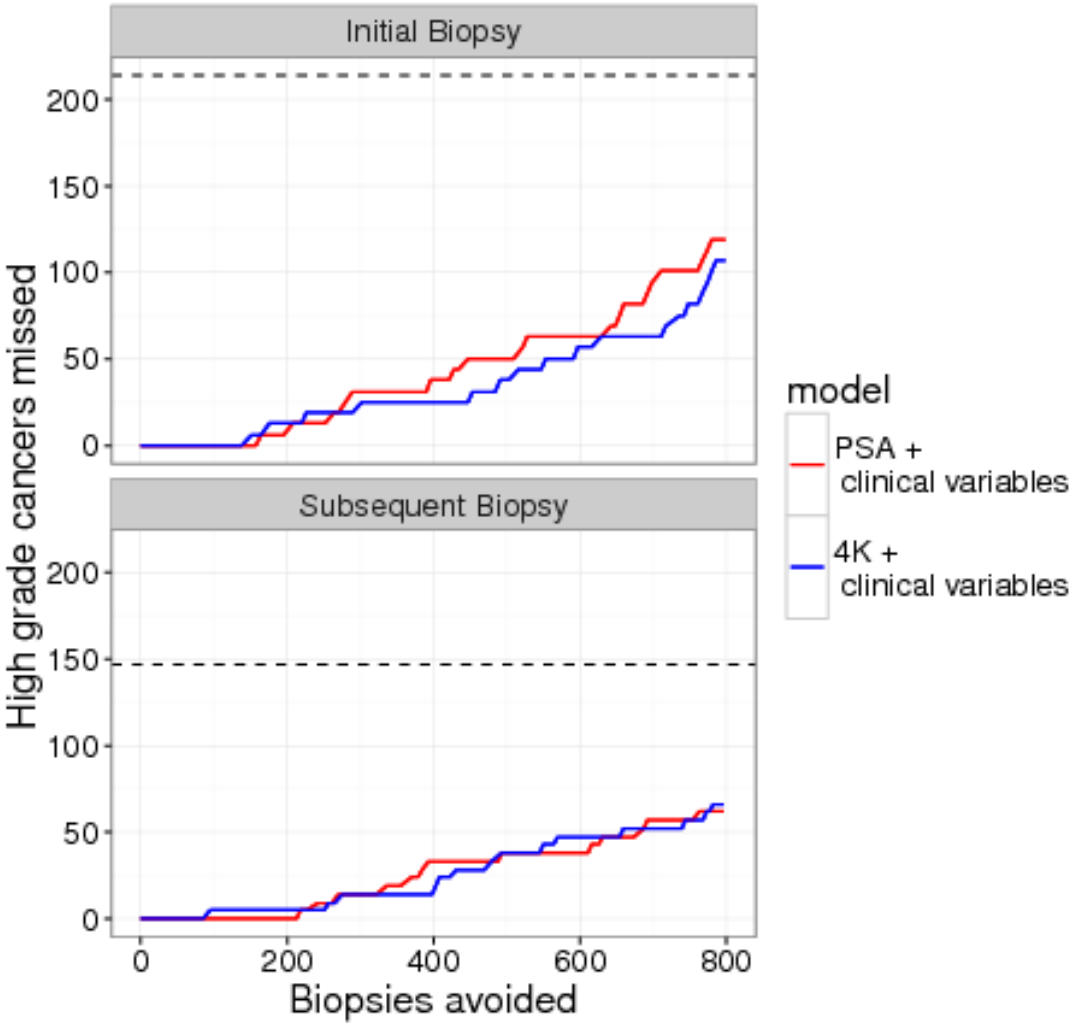


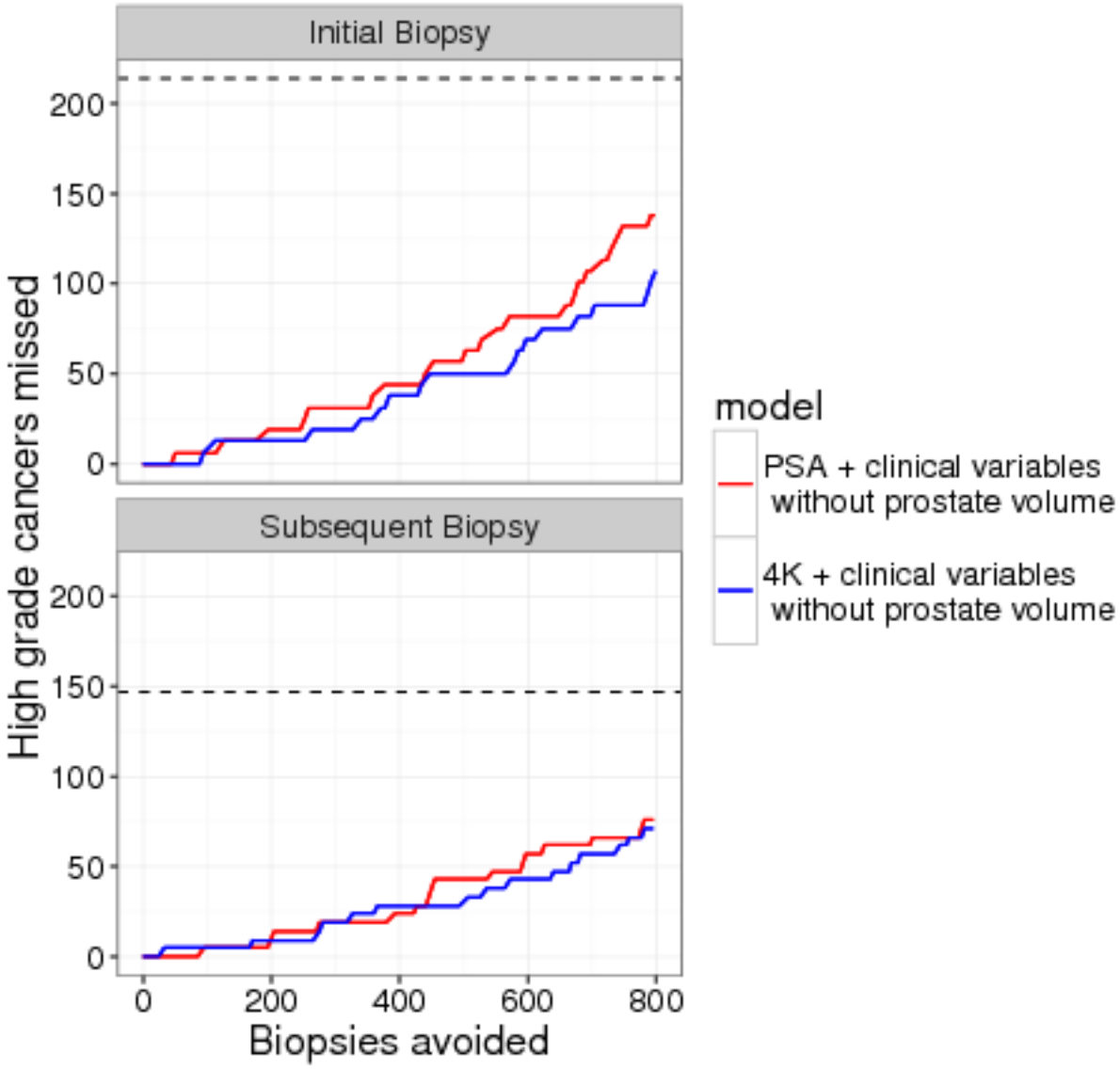

Initial Biopsy
Subsequent Biopsy
sensitivity
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
1−specificity
model
PSA + clinical variables without prostate volume
4K + clinical variables without prostate volume

**Supp Figure 3.**



**Sup Figure 4.**


Initial Biopsy
Subsequent Biopsy
High grade cancers missed
200
150
100
50
0
0
200
400
600
800
Biopsies avoided
model
PSA + clinical variables
4K + clinical variables


Initial Biopsy
Subsequent Biopsy
High grade cancers missed
Biopsies avoided
0
50
100
150
200
0
200
400
600
800
model
PSA + clinical variables without prostate volume
4K + clinical variables without prostate volume

**Supplementary Material Figure and Table legends**

**Supp Table 1.** Biopsy characteristics at each sequential surveillance biopsy in 339 participants in the test set. Medians with interquartile range (IQR) or numbers with percent (%) are shown.

**Supp Table 2:** Summary of fitted models using clinical variables without prostate volume + PSA or 4K panel.

**Supp Figure 1.** Calibration assessment for models with prostate volume (top) and without prostate volume (bottom).

**Supp Figure 2.** ROC Curves for full models (top panels) and models without prostate volume (bottom panels).

**Supp Figure 3.** Decision Curve Analysis for models without prostate volume.

**Sup Figure 4.** Lorenz curves for models with (top) and without (bottom) prostate volume. Number of biopsies avoided by high grade cancers missed for 1000 men in each group. Dashed black line represents total number of expected high grade cancers per 1000 men. Results are shown for high risk thresholds of 0-0.30.